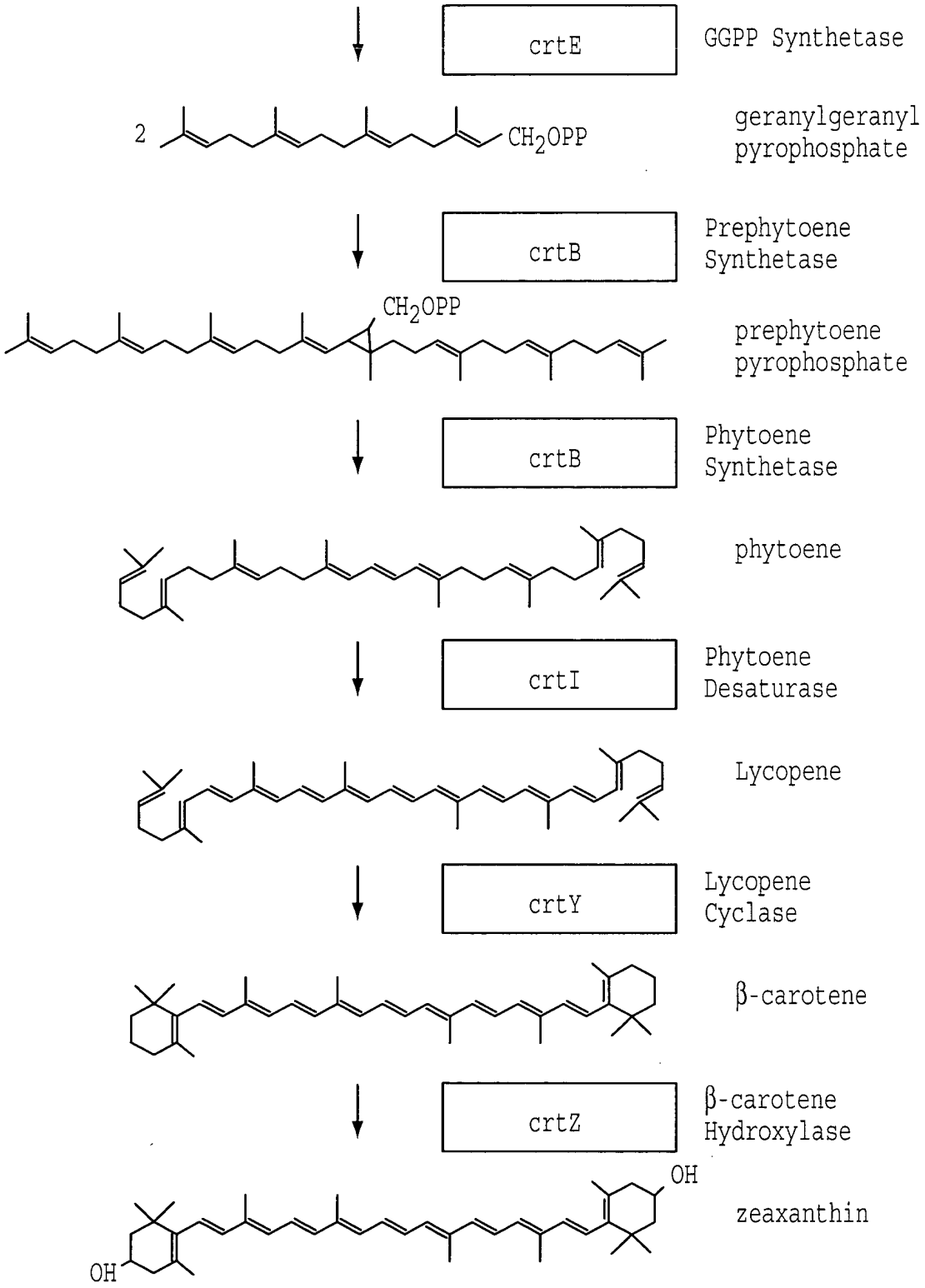

crtE
GGPP Synthetase
2
CH_2OPP
geranylgeranyl pyrophosphate
crtB
Prephytoene Synthetase
CH_2OPP
prephytoene pyrophosphate
crtB
Phytoene Synthetase
phytoene
crtI
Phytoene Desaturase
Lycopene
crtY
Lycopene Cyclase
β-carotene
crtZ
β-carotene Hydroxylase
OH
OH
zeaxanthin

FIG. 1

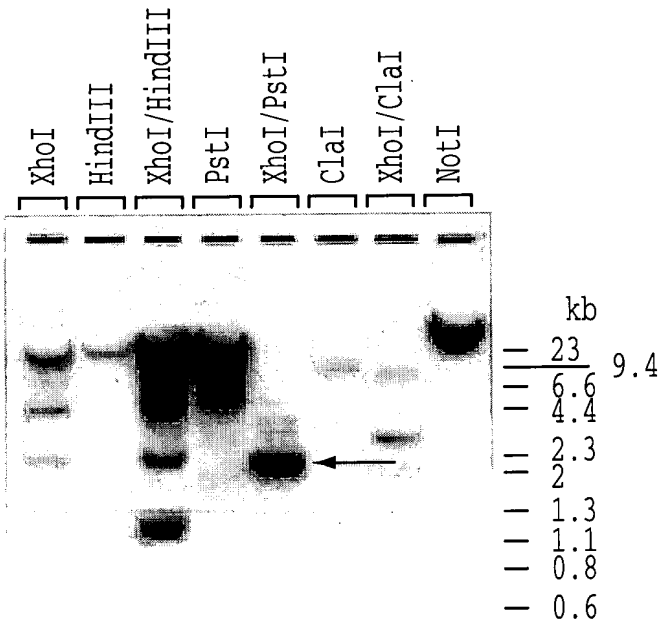

XhoI
HindIII
XhoI/HindIII
PstI
XhoI/PstI
ClaI
XhoI/ClaI
NotI
kb
23
9.4
6.6
4.4
2.3
2
1.3
1.1
0.8
0.6

FIG. 2

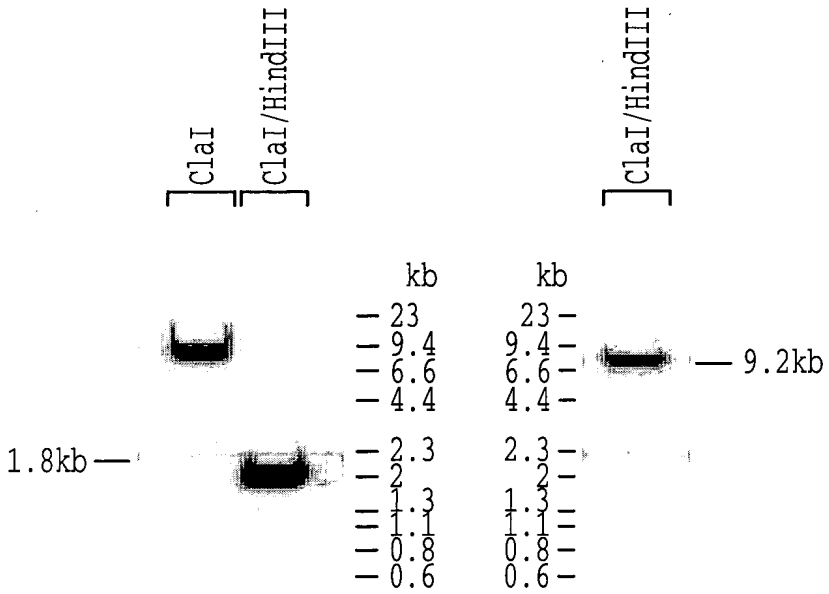

ClaI
ClaI/HindIII
kb
23
9.4
6.6
4.4
2.3
2
1.3
1.1
0.8
0.6
1.8kb

FIG. 3A

ClaI/HindIII
kb
23
9.4
6.6
4.4
2.3
2
1.3
1.1
0.8
0.6
9.2kb

FIG. 3B

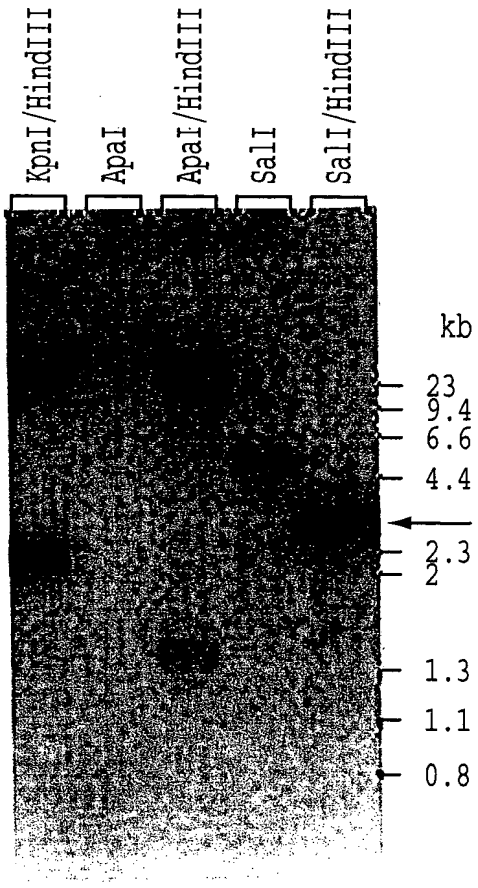

KpnI/HindIII
ApaI
ApaI/HindIII
SalI
SalI/HindIII
kb
23
9.4
6.6
4.4
2.3
2
1.3
1.1
0.8

FIG. 4

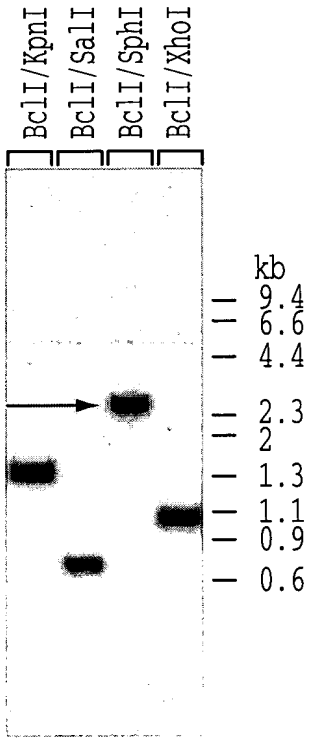

BclI/KpnI
BclI/SalI
BclI/SphI
BclI/XhoI
kb
9.4
6.6
4.4
2.3
2
1.3
1.1
0.9
0.6

FIG. 5

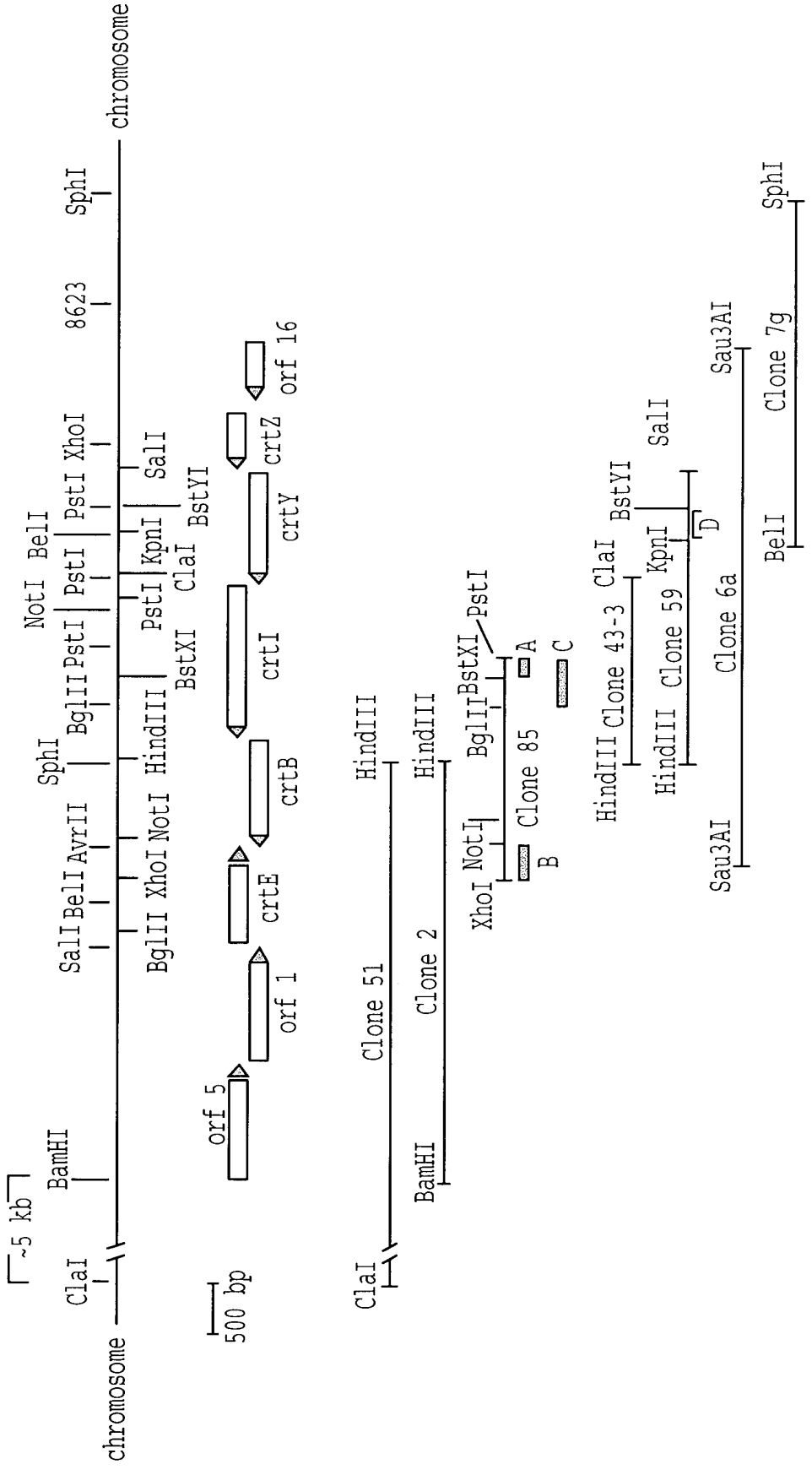

chromosome
SphI
8623
NotI BelI
PstI XhoI
PstI
KpnI SalI
ClaI BstYI
BstXI
PstI
SphI
BglII PstI
HindIII
crtZ
orf 16
crtY
crtI
crtB
crtE
orf 1
orf 5
SalI BelI AvrII
BglII XhoI NotI
BamHI
~5 kb
ClaI
500 bp
chromosome
HindIII
Clone 51
BamHI
HindIII
Clone 2
XhoI NotI
BglII BstXI PstI
Clone 85
A
B
C
HindIII Clone 43-3
ClaI
BstYI
HindIII Clone 59
KpnI
SalI
D
Sau3AI
Clone 6a
Sau3AI
BelI
Clone 7g
SphI

FIG. 6

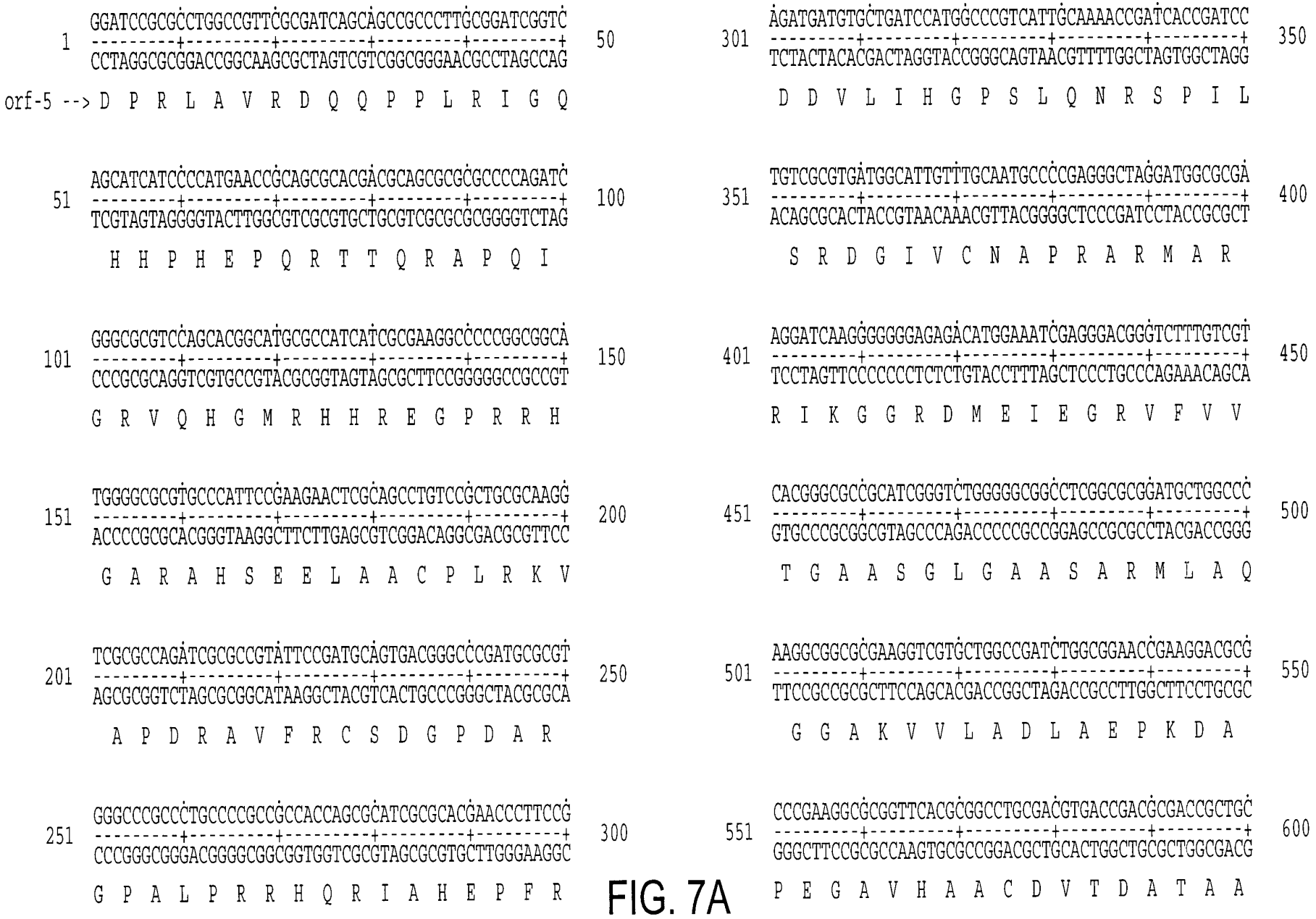

1 GGATCCGCGCCTGGCCGTTCGCGATCAGCAGCCGCCCTTGCGGATCGGTC 50
CCTAGGCGCGGACCGGCAAGCGCTAGTCGTCGGCGGGAACGCCTAGCCAG
orf-5 --> D P R L A V R D Q Q P P L R I G Q
51 AGCATCATCCCCATGAACCGCAGCGCACGACGCAGCGCGCGCCCCAGATC 100
TCGTAGTAGGGGTACTTGGCGTCGCGTGCTGCGTCGCGCGCGGGGTCTAG
H H P H E P Q R T T Q R A P Q I
101 GGGCGCGTCCAGCACGGCATGCGCCATCATCGCGAAGGCCCCCGGCGGCA 150
CCCGCGCAGGTCGTGCCGTACGCGGTAGTAGCGCTTCCGGGGGCCGCCGT
G R V Q H G M R H H R E G P R R H
151 TGGGGCGCGTGCCCATTCCGAAGAACTCGCAGCCTGTCCGCTGCGCAAGG 200
ACCCCGCGCACGGGTAAGGCTTCTTGAGCGTCGGACAGGCGACGCGTTCC
G A R A H S E E L A A C P L R K V
201 TCGCGCCAGATCGCGCCGTATTCCGATGCAGTGACGGGCCCGATGCGCGT 250
AGCGCGGTCTAGCGCGGCATAAGGCTACGTCACTGCCCGGGCTACGCGCA
A P D R A V F R C S D G P D A R
251 GGGCCCGCCCTGCCCCGCCGCCACCAGCGCATCGCGCACGAACCCTTCCG 300
CCCGGGCGGGACGGGGCGGCGGTGGTCGCGTAGCGCGTGCTTGGGAAGGC
G P A L P R R H Q R I A H E P F R
301 AGATGATGTGCTGATCCATGGCCCGTCATTGCAAAACCGATCACCGATCC 350
TCTACTACACGACTAGGTACCGGGCAGTAACGTTTTGGCTAGTGGCTAGG
D D V L I H G P S L Q N R S P I L
351 TGTCGCGTGATGGCATTGTTTGCAATGCCCCGAGGGCTAGGATGGCGCGA 400
ACAGCGCACTACCGTAACAAACGTTACGGGGCTCCCGATCCTACCGCGCT
S R D G I V C N A P R A R M A R
401 AGGATCAAGGGGGGGAGAGACATGGAAATCGAGGGACGGGTCTTTGTCGT 450
TCCTAGTTCCCCCCCTCTCTGTACCTTTAGCTCCCTGCCCAGAAACAGCA
R I K G G R D M E I E G R V F V V
451 CACGGGCGCCGCATCGGGTCTGGGGGCGGCCTCGGCGCGGATGCTGGCCC 500
GTGCCCGCGGCGTAGCCCAGACCCCCGCCGGAGCCGCGCCTACGACCGGG
T G A A S G L G A A S A R M L A Q
501 AAGGCGGCGCGAAGGTCGTGCTGGCCGATCTGGCGGAACCGAAGGACGCG 550
TTCCGCCGCGCTTCCAGCACGACCGGCTAGACCGCCTTGGCTTCCTGCGC
G G A K V V L A D L A E P K D A
551 CCCGAAGGCGCGGTTCACGCGGCCTGCGACGTGACCGACGCGACCGCTGC 600
GGGCTTCCGCGCCAAGTGCGCCGGACGCTGCACTGGCTGCGCTGGCGACG
P E G A V H A A C D V T D A T A A

FIG. 7A

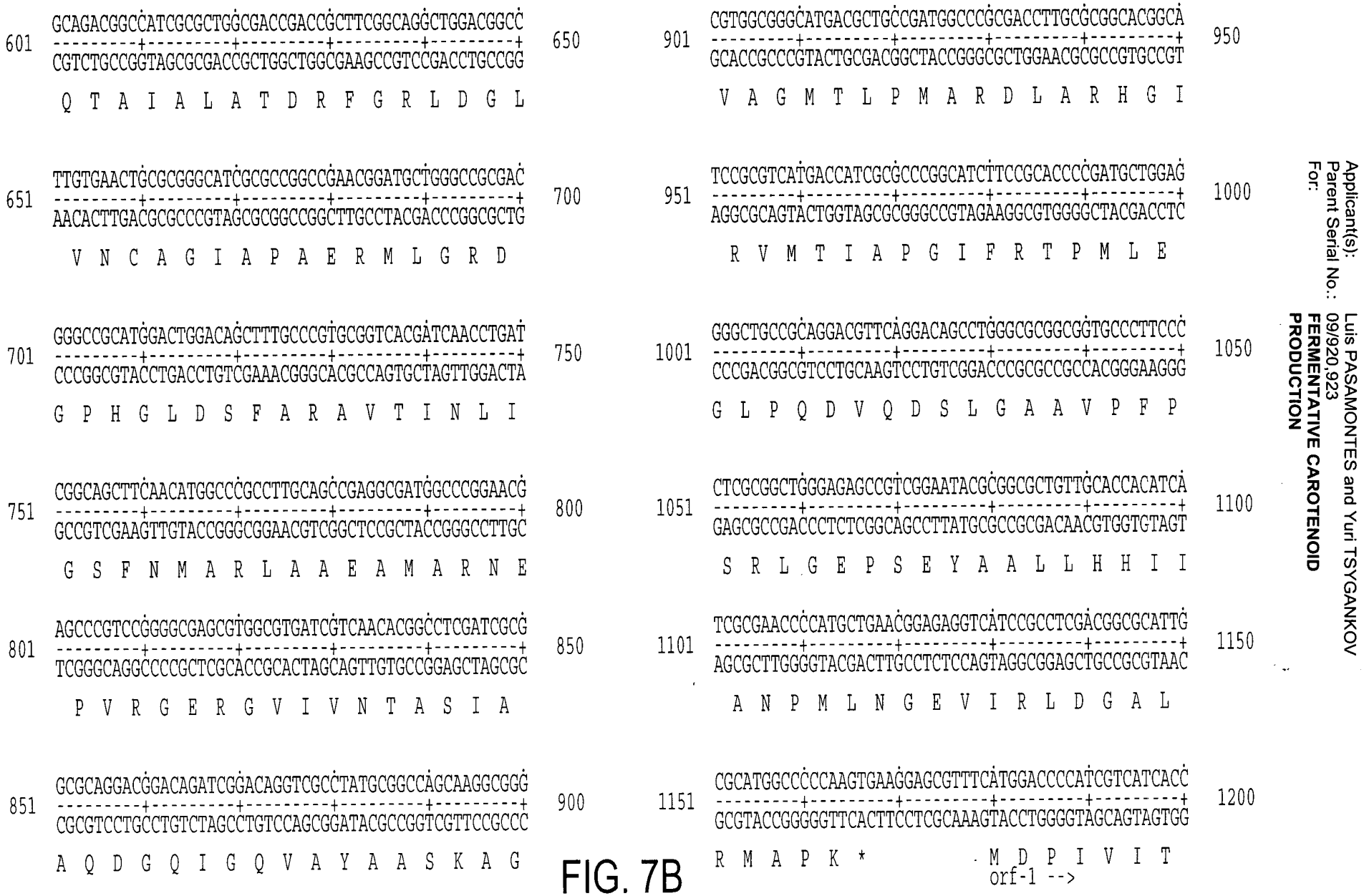

601 GCAGACGGCCATCGCGCTGGCGACCGACCGCTTCGGCAGGCTGGACGGCC 650
CGTCTGCCGGTAGCGCGACCGCTGGCTGGCGAAGCCGTCCGACCTGCCGG
Q T A I A L A T D R F G R L D G L
651 TTGTGAACTGCGCGGGCATCGCGCCGGCCGAACGGATGCTGGGCCGCGAC 700
AACACTTGACGCGCCCGTAGCGCGGCCGGCTTGCCTACGACCCGGCGCTG
V N C A G I A P A E R M L G R D
701 GGGCCGCATGGACTGGACAGCTTTGCCCGTGCGGTCACGATCAACCTGAT 750
CCCGGCGTACCTGACCTGTCGAAACGGGCACGCCAGTGCTAGTTGGACTA
G P H G L D S F A R A V T I N L I
751 CGGCAGCTTCAACATGGCCCGCCTTGCAGCCGAGGCGATGGCCCGGAACG 800
GCCGTCGAAGTTGTACCGGGCGGAACGTCGGCTCCGCTACCGGGCCTTGC
G S F N M A R L A A E A M A R N E
801 AGCCCGTCCGGGGCGAGCGTGGCGTGATCGTCAACACGGCCTCGATCGCG 850
TCGGGCAGGCCCCGCTCGCACCGCACTAGCAGTTGTGCCGGAGCTAGCGC
P V R G E R G V I V N T A S I A
851 GCGCAGGACGGACAGATCGGACAGGTCGCCTATGCGGCCAGCAAGGCGGG 900
CGCGTCCTGCCTGTCTAGCCTGTCCAGCGGATACGCCGGTCGTTCCGCCC
A Q D G Q I G Q V A Y A A S K A G
901 CGTGGCGGGCATGACGCTGCCGATGGCCCGCGACCTTGCGCGGCACGGCA 950
GCACCGCCCGTACTGCGACGGCTACCGGGCGCTGGAACGCGCCGTGCCGT
V A G M T L P M A R D L A R H G I
951 TCCGCGTCATGACCATCGCGCCCGGCATCTTCCGCACCCCGATGCTGGAG 1000
AGGCGCAGTACTGGTAGCGCGGGCCGTAGAAGGCGTGGGGCTACGACCTC
R V M T I A P G I F R T P M L E
1001 GGGCTGCCGCAGGACGTTCAGGACAGCCTGGGCGCGGCGGTGCCCTTCCC 1050
CCCGACGGCGTCCTGCAAGTCCTGTCGGACCCGCGCCGCCACGGGAAGGG
G L P Q D V Q D S L G A A V P F P
1051 CTCGCGGCTGGGAGAGCCGTCGGAATACGCGGCGCTGTTGCACCACATCA 1100
GAGCGCCGACCCTCTCGGCAGCCTTATGCGCCGCGACAACGTGGTGTAGT
S R L G E P S E Y A A L L H H I I
1101 TCGCGAACCCCATGCTGAACGGAGAGGTCATCCGCCTCGACGGCGCATTG 1150
AGCGCTTGGGGTACGACTTGCCTCTCCAGTAGGCGGAGCTGCCGCGTAAC
A N P M L N G E V I R L D G A L
1151 CGCATGGCCCCCAAGTGAAGGAGCGTTTCATGGACCCCATCGTCATCACC 1200
GCGTACCGGGGGGTTCACTTCCTCGCAAAGTACCTGGGGTAGCAGTAGTGG
R M A P K * M D P I V I T
orf-1 -->

FIG. 7B

```
1201 GGCGCGATGCGCACCCCGATGGGGGCATTCCAGGGCGATCTTGCCGCGAT 1250
     CCGCGCTACGCGTGGGGCTACCCCCGTAAGGTCCCGCTAGAACGGCGCTA
     G  A  M  R  T  P  M  G  A  F  Q  G  D  L  A  A  M

1251 GGATGCCCCGACCCTTGGCGCGGACGCGATCCGCGCCGCGCTGAACGGCC 1300
     CCTACGGGGCTGGGAACCGCGCCTGCGCTAGGCGCGGCGCGACTTGCCGG
     D  A  P  T  L  G  A  D  A  I  R  A  A  L  N  G  L

1301 TGTCGCCCGACATGGTGGACGAGGTGCTGATGGGCTGCGTCCTCGCCGCG 1350
     ACAGCGGGCTGTACCACCTGCTCCACGACTACCCGACGCAGGAGCGGCGC
     S  P  D  M  V  D  E  V  L  M  G  C  V  L  A  A

1351 GGCCAGGGTCAGGCACCGGCACGTCAGGCGGCGCTTGGCGCCGGACTGCC 1400
     CCGGTCCCAGTCCGTGGCCGTGCAGTCCGCCGCGAACCGCGGCCTGACGG
     G  Q  G  Q  A  P  A  R  Q  A  A  L  G  A  G  L  P

1401 GCTGTCGACGGGCACGACCACCATCAACGAGATGTGCGGATCGGGCATGA 1450
     CGACAGCTGCCCGTGCTGGTGGTAGTTGCTCTACACGCCTAGCCCGTACT
     L  S  T  G  T  T  T  I  N  E  M  C  G  S  G  M  K

1451 AGGCCGCGATGCTGGGCCATGACCTGATCGCCGCGGGATCGGCGGGCATC 1500
     TCCGGCGCTACGACCCGGTACTGGACTAGCGGCGCCCTAGCCGCCCGTAG
     A  A  M  L  G  H  D  L  I  A  A  G  S  A  G  I

1501 GTCGTCGCCGGCGGGATGGAGAGCATGTCGAACGCCCCCTACCTGCTGCC 1550
     CAGCAGCGGCCGCCCTACCTCTCGTACAGCTTGCGGGGGATGGACGACGG
     V  V  A  G  G  M  E  S  M  S  N  A  P  Y  L  L  P

1551 CAAGGCGCGGTCGGGGATGCGCATGGGCCATGACCGTGTGCTGGATCACA 1600
     GTTCCGCGCCAGCCCCTACGCGTACCCGGTACTGGCACACGACCTAGTGT
     K  A  R  S  G  M  R  M  G  H  D  R  V  L  D  H  M

1601 TGTTCCTCGACGGGTTGGAGGACGCCTATGACAAGGGCCGCCTGATGGGC 1650
     ACAAGGAGCTGCCCAACCTCCTGCGGATACTGTTCCCGGCGGACTACCCG
     F  L  D  G  L  E  D  A  Y  D  K  G  R  L  M  G

1651 ACCTTCGCCGAGGATTGCGCCGGCGATCACGGTTTCACCCGCGAGGCGCA 1700
     TGGAAGCGGCTCCTAACGCGGCCGCTAGTGCCAAAGTGGGCGCTCCGCGT
     T  F  A  E  D  C  A  G  D  H  G  F  T  R  E  A  Q

1701 GGACGACTATGCGCTGACCAGCCTGGCCCGCGCGCAGGACGCCATCGCCA 1750
     CCTGCTGATACGCGACTGGTCGGACCGGGCGCGCGTCCTGCGGTAGCGGT
     D  D  Y  A  L  T  S  L  A  R  A  Q  D  A  I  A  S

1751 GCGGTGCCTTCGCCGCCGAGATCGCGCCCGTGACCGTCACGGCACGCAAG 1800
     CGCCACGGAAGCGGCGGCTCTAGCGCGGGCACTGGCAGTGCCGTGCGTTC
     G  A  F  A  A  E  I  A  P  V  T  V  T  A  R  K
```

FIG. 7C

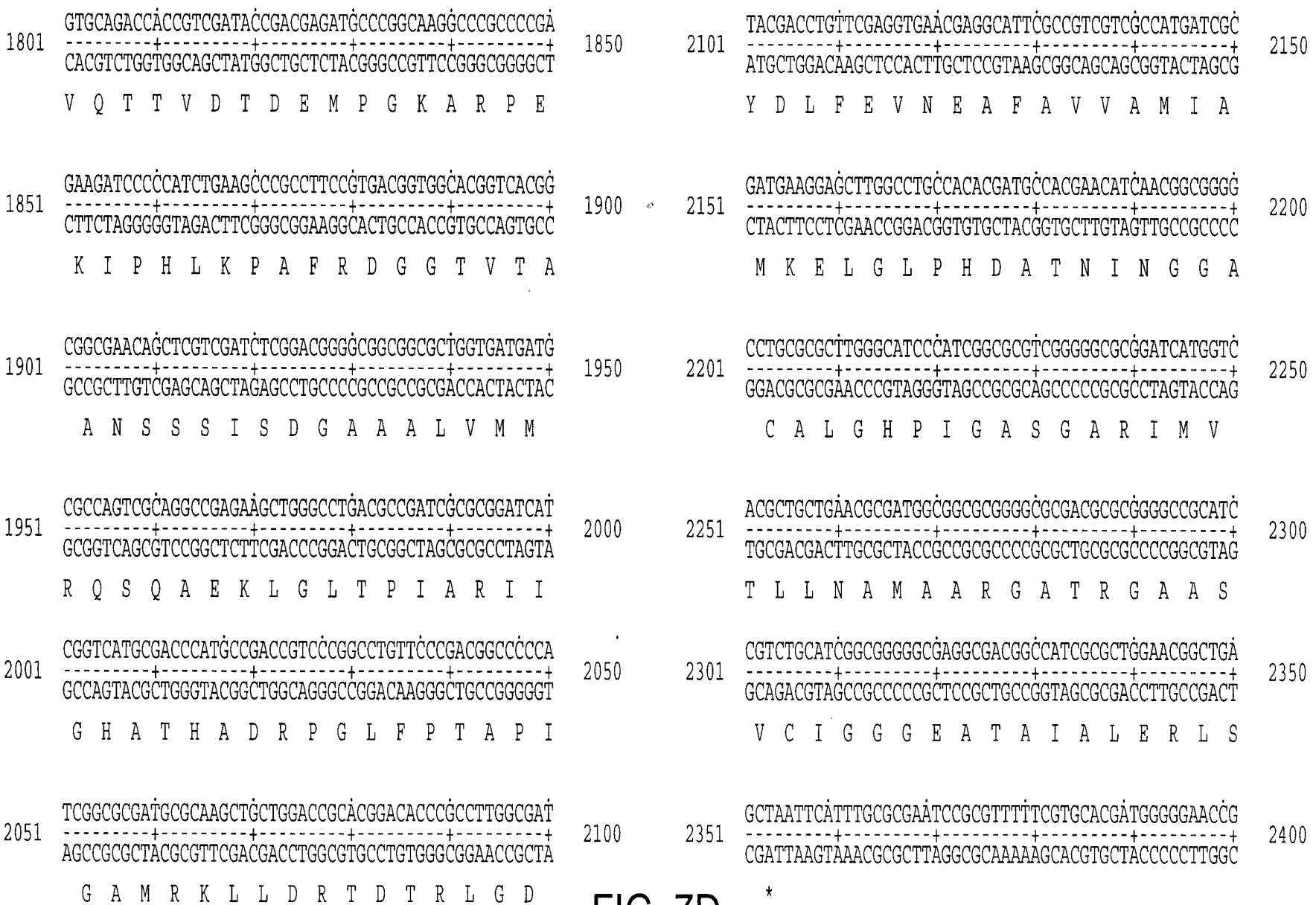

1801 GTGCAGACCACCGTCGATACCGACGAGATGCCCGGCAAGGCCCGCCCCCGA 1850
CACGTCTGGTGGCAGCTATGGCTGCTCTACGGGCCGTTCCGGGCGGGGGCT
V Q T T V D T D E M P G K A R P E
1851 GAAGATCCCCCATCTGAAGCCCGCCTTCCGTGACGGTGGCACGGTCACGG 1900
CTTCTAGGGGGTAGACTTCGGGCGGAAGGCACTGCCACCGTGCCAGTGCC
K I P H L K P A F R D G G T V T A
1901 CGGCGAACAGCTCGTCGATCTCGGACGGGGCGGCGGCGCTGGTGATGATG 1950
GCCGCTTGTCGAGCAGCTAGAGCCTGCCCCGCCGCCGCGACCACTACTAC
A N S S S I S D G A A A L V M M
1951 CGCCAGTCGCAGGCCGAGAAGCTGGGCCTGACGCCGATCGCGCGGATCAT 2000
GCGGTCAGCGTCCGGCTCTTCGACCCGGACTGCGGCTAGCGCGCCTAGTA
R Q S Q A E K L G L T P I A R I I
2001 CGGTCATGCGACCCATGCCGACCGTCCCGGCCTGTTCCCGACGGCCCCCA 2050
GCCAGTACGCTGGGTACGGCTGGCAGGGCCGGACAAGGGCTGCCGGGGGT
G H A T H A D R P G L F P T A P I
2051 TCGGCGCGATGCGCAAGCTGCTGGACCGCACGGACACCCGCCTTGGCGAT 2100
AGCCGCGCTACGCGTTCGACGACCTGGCGTGCCTGTGGGCGGAACCGCTA
G A M R K L L D R T D T R L G D
2101 TACGACCTGTTCGAGGTGAACGAGGCATTCGCCGTCGTCGCCATGATCGC 2150
ATGCTGGACAAGCTCCACTTGCTCCGTAAGCGGCAGCAGCGGTACTAGCG
Y D L F E V N E A F A V V A M I A
2151 GATGAAGGAGCTTGGCCTGCCACACGATGCCACGAACATCAACGGCGGGG 2200
CTACTTCCTCGAACCGGACGGTGTGCTACGGTGCTTGTAGTTGCCGCCCC
M K E L G L P H D A T N I N G G A
2201 CCTGCGCGCTTGGGCATCCCATCGGCGCGTCGGGGGCGCGGATCATGGTC 2250
GGACGCGCGAACCCGTAGGGTAGCCGCGCAGCCCCCGCGCCTAGTACCAG
C A L G H P I G A S G A R I M V
2251 ACGCTGCTGAACGCGATGGCGGCGCGGGGCGCGACGCGCGGGGCCGCATC 2300
TGCGACGACTTGCGCTACCGCCGCGCCCCGCGCTGCGCGCCCCGGCGTAG
T L L N A M A A R G A T R G A A S
2301 CGTCTGCATCGGCGGGGGCGAGGCGACGGCCATCGCGCTGGAACGGCTGA 2350
GCAGACGTAGCCGCCCCCGCTCCGCTGCCGGTAGCGCGACCTTGCCGACT
V C I G G G E A T A I A L E R L S
2351 GCTAATTCATTTGCGCGAATCCGCGTTTTTCGTGCACGATGGGGGAACCG 2400
CGATTAAGTAAACGCGCTTAGGCGCAAAAAGCACGTGCTACCCCCTTGGC
*

FIG. 7D

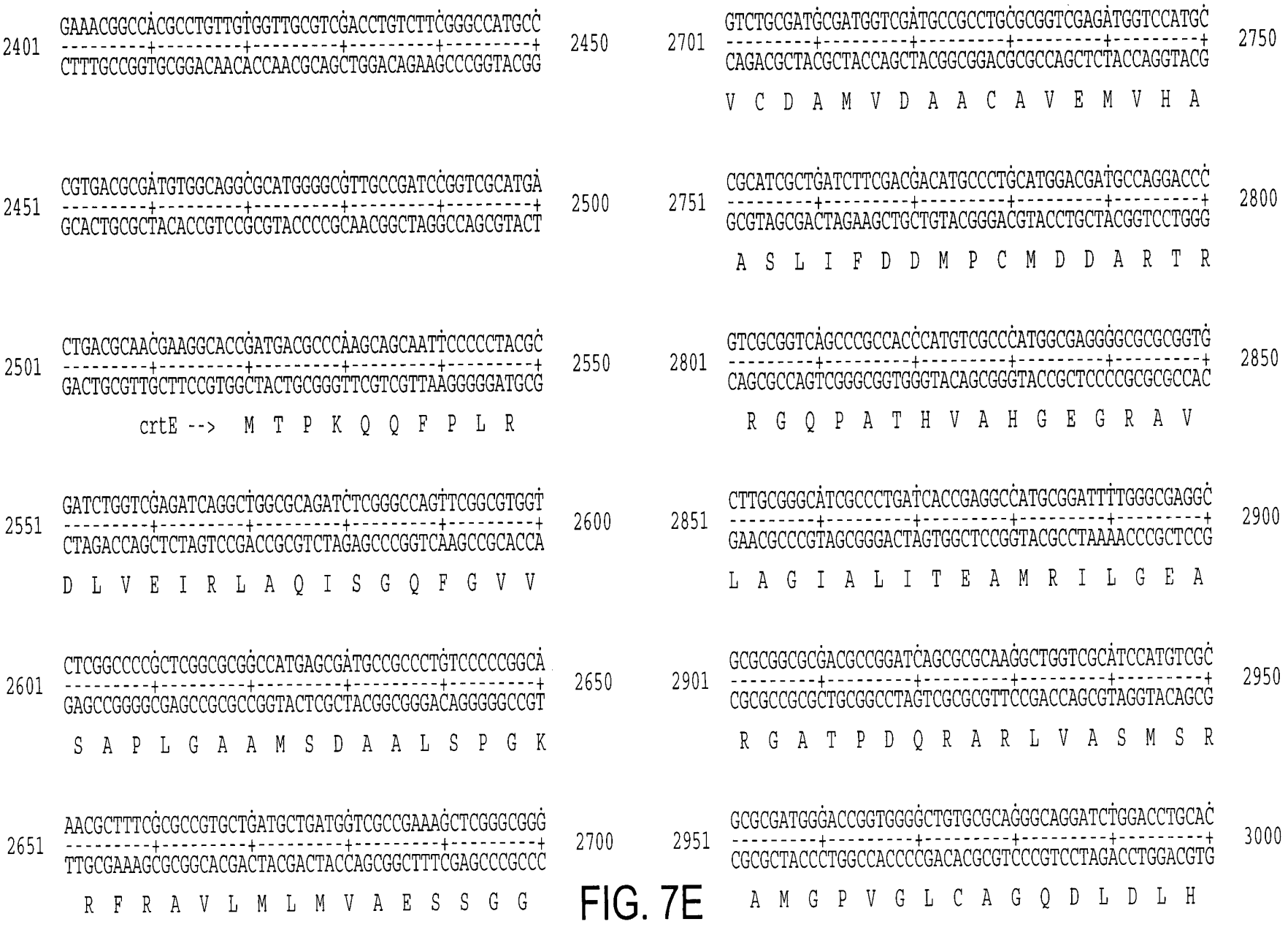
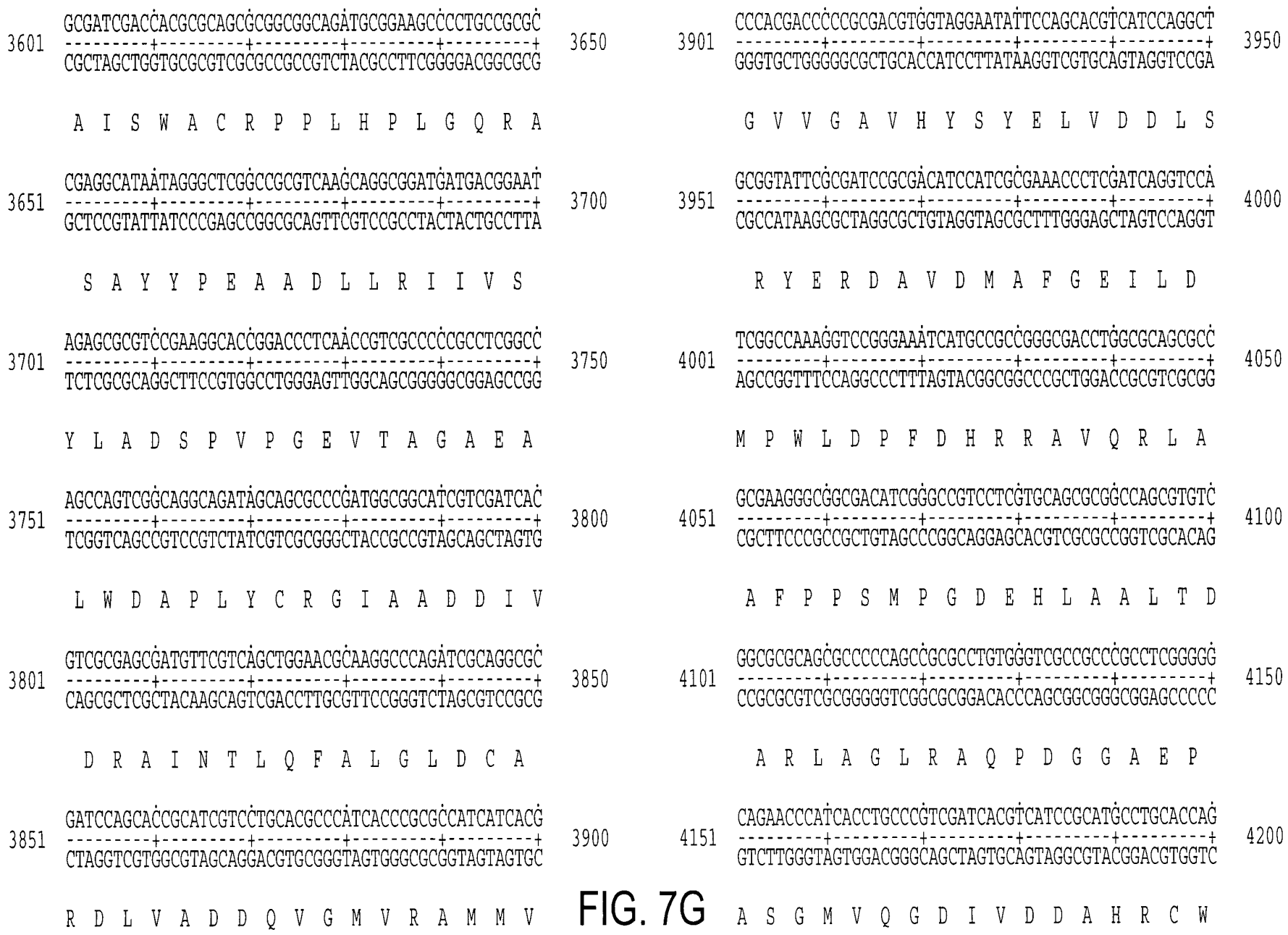
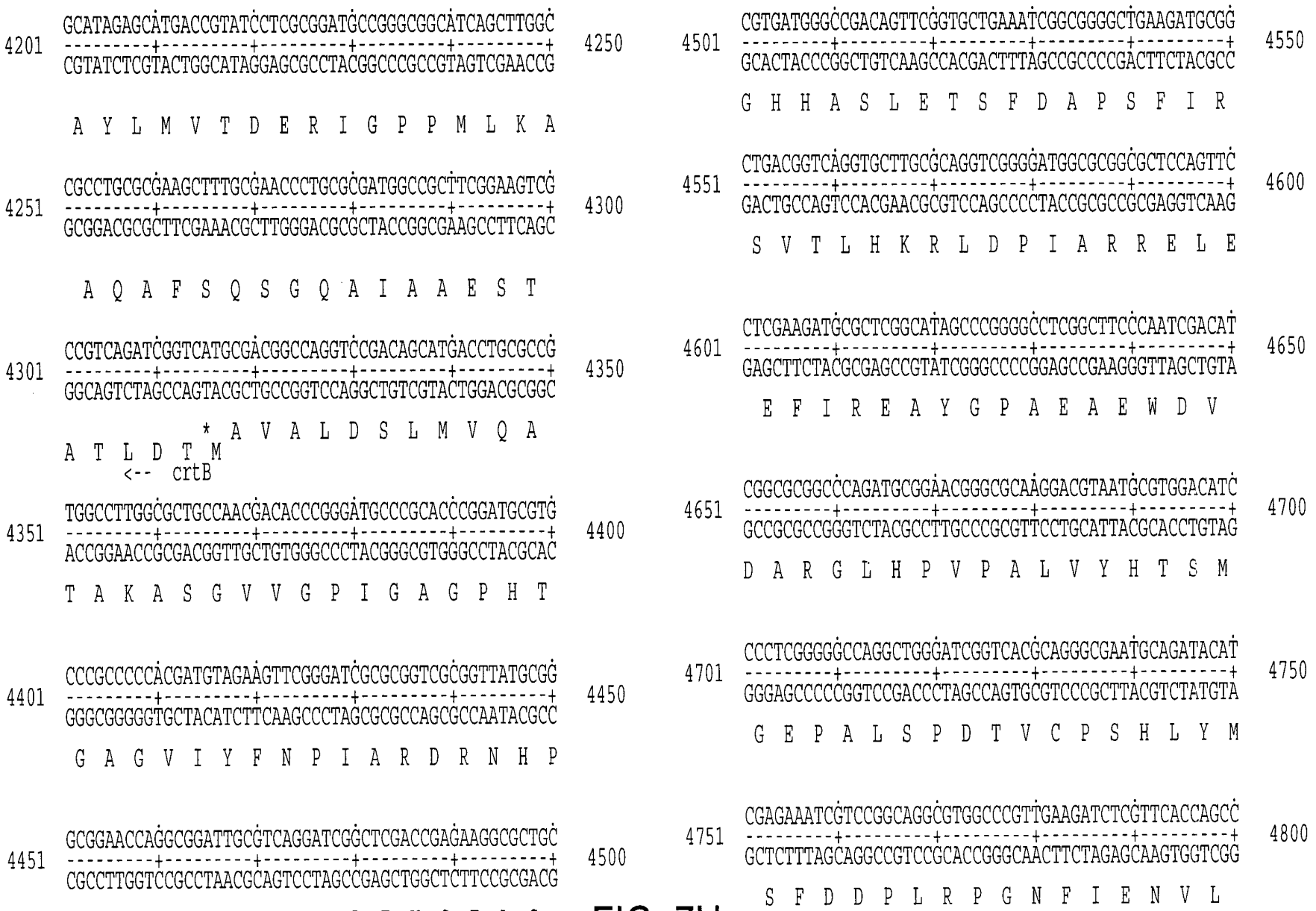

2401 GAAACGGCCACGCCTGTTGTGGTTGCGTCGACCTGTCTTCGGGCCATGCC
CTTTGCCGGTGCGGACAACACCAACGCAGCTGGACAGAAGCCCGGTACGG
2450
2501 CTGACGCAACGAAGGCACCGATGACGCCCAAGCAGCAATTCCCCCTACGC
GACTGCGTTGCTTCCGTGGCTACTGCGGGTTCGTCGTTAAGGGGGATGCG
crtE --> M T P K Q Q F P L R
2550
2551 GATCTGGTCGAGATCAGGCTGGCGCAGATCTCGGGCCAGTTCGGCGTGGT
CTAGACCAGCTCTAGTCCGACCGCGTCTAGAGCCCGGTCAAGCCGCACCA
D L V E I R L A Q I S G Q F G V V
2600
2601 CTCGGCCCCGCTCGGCGCGGCCATGAGCGATGCCGCCCTGTCCCCCGGCA
GAGCCGGGGCGAGCCGCGCCGGTACTCGCTACGGCGGGACAGGGGGCCGT
S A P L G A A M S D A A L S P G K
2650
2651 AACGCTTTCGCGCCGTGCTGATGCTGATGGTCGCCGAAAGCTCGGGCGGG
TTGCGAAAGCGCGGCACGACTACGACTACCAGCGGCTTTCGAGCCCGCCC
R F R A V L M L M V A E S S G G
2700
2701 GTCTGCGATGCGATGGTCGATGCCGCCTGCGCGGTCGAGATGGTCCATGC
CAGACGCTACGCTACCAGCTACGGCGGACGCGCCAGCTCTACCAGGTACG
V C D A M V D A A C A V E M V H A
2750
2751 CGCATCGCTGATCTTCGACGACATGCCCTGCATGGACGATGCCAGGACCC
GCGTAGCGACTAGAAGCTGCTGTACGGGACGTACCTGCTACGGTCCTGGG
A S L I F D D M P C M D D A R T R
2800
2801 GTCGCGGTCAGCCCGCCACCCATGTCGCCCATGGCGAGGGGCGCGCGGTG
CAGCGCCAGTCGGGCGGTGGGTACAGCGGGTACCGCTCCCCGCGCGCCAC
R G Q P A T H V A H G E G R A V
2850
2851 CTTGCGGGCATCGCCCTGATCACCGAGGCCATGCGGATTTTGGGCGAGGC
GAACGCCCGTAGCGGGACTAGTGGCTCCGGTACGCCTAAAACCCGCTCCG
L A G I A L I T E A M R I L G E A
2900
2901 GCGCGGCGCGACGCCGGATCAGCGCGCAAGGCTGGTCGCATCCATGTCGC
CGCGCCGCGCTGCGGCCTAGTCGCGCGTTCCGACCAGCGTAGGTACAGCG
R G A T P D Q R A R L V A S M S R
2950
2951 GCGCGATGGGACCGGTGGGGCTGTGCGCAGGGCAGGATCTGGACCTGCAC
CGCGCTACCCTGGCCACCCCGACACGCGTCCCGTCCTAGACCTGGACGTG
A M G P V G L C A G Q D L D L H
3000

FIG. 7E

```
3001 GCCCCCAAGGACGCCGCCGGGATCGAACGTGAACAGGACCTCAAGACCGG 3050
     CGGGGGTTCCTGCGGCGGCCCTAGCTTGCACTTGTCCTGGAGTTCTGGCC
       A  P  K  D  A  A  G  I  E  R  E  Q  D  L  K  T  G

3051 CGTGCTGTTCGTCGCGGGCCTCGAGATGCTGTCCATTATTAAGGGTCTGG 3100
     GCACGACAAGCAGCGCCCGGAGCTCTACGACAGGTAATAATTCCCAGACC
       V  L  F  V  A  G  L  E  M  L  S  I  I  K  G  L  D

3101 ACAAGGCCGAGACCGAGCAGCTCATGGCCTTCGGGCGTCAGCTTGGTCGG 3150
     TGTTCCGGCTCTGGCTCGTCGAGTACCGGAAGCCCGCAGTCGAACCAGCC
       K  A  E  T  E  Q  L  M  A  F  G  R  Q  L  G  R

3151 GTCTTCCAGTCCTATGACGACCTGCTGGACGTGATCGGCGACAAGGCCAG 3200
     CAGAAGGTCAGGATACTGCTGGACGACCTGCACTAGCCGCTGTTCCGGTC
       V  F  Q  S  Y  D  D  L  L  D  V  I  G  D  K  A  S

3201 CACCGGCAAGGATACGGCGCGCGACACCGCCGCCCCCGGCCCAAAGGGCG 3250
     GTGGCCGTTCCTATGCCGCGCGCTGTGGCGGCGGGGGCCGGGTTTCCCGC
       T  G  K  D  T  A  R  D  T  A  A  P  G  P  K  G  G

3251 GCCTGATGGCGGTCGGACAGATGGGCGACGTGGCGCAGCATTACCGCGCC 3300
     CGGACTACCGCCAGCCTGTCTACCCGCTGCACCGCGTCGTAATGGCGCGG
       L  M  A  V  G  Q  M  G  D  V  A  Q  H  Y  R  A

3301 AGCCGCGCGCAACTGGACGAGCTGATGCGCACCCGGCTGTTCCGCGGGGG 3350
     TCGGCGCGCGTTGACCTGCTCGACTACGCGTGGGCCGACAAGGCGCCCCC
       S  R  A  Q  L  D  E  L  M  R  T  R  L  F  R  G  G

3351 GCAGATCGCGGACCTGCTGGCCCGCGTGCTGCCGCATGACATCCGCCGCA 3400
     CGTCTAGCGCCTGGACGACCGGGCGCACGACGGCGTACTGTAGGCGGCGT
       Q  I  A  D  L  L  A  R  V  L  P  H  D  I  R  R  S

3401 GCGCCTAGGCGCGCGGTCGGGTCCACAGGCCGTCGCGGCTGATTTCGCCG 3450
     CGCGGATCCGCGCGCCAGCCCAGGTGTCCGGCAGCGCCGACTAAAGCGGC
       A  *
          *  A  R  P  R  T  W  L  G  D  R  S  I  E  G

3451 CCGCGCAGGCGCGATGCGGCCGCGTCCAAGCCTCCGCGCGCCAGAAGCCC 3500
     GGCGCGTCCGCGCTACGCCGGCGCAGGTTCGGAGGCGCGCGGTCTTCGGG
       G  R  L  R  S  A  A  A  D  L  G  G  R  A  L  L  G

3501 GATCTTGGCAGCCTTCGACGTGCTGATCCGCTGGCGATAGGCCTCGGGGC 3550
     CTAGAACCGTCGGAAGCTGCACGACTAGGCGACCGCTATCCGGAGCCCCG
       I  K  A  A  K  S  T  S  I  R  Q  R  Y  A  E  P

3551 CACCCTGCCGGATGCGCGTCCCGATTGCGCGATAGATACGCAGCGCGGCG 3600
     GTGGGACGGCCTACGCGCAGGGCTAACGCGCTATCTATGCGTCGCGCCGC
       G  G  Q  R  I  R  T  G  I  A  R  Y  I  R  L  A  A
```

FIG. 7F


3601 GCGATCGACCACGCGCAGCGCGGCGGCAGATGCGGAAGCCCCTGCCGCGC 3650
CGCTAGCTGGTGCGCGTCGCGCCGCCGTCTACGCCTTCGGGGACGGCGCG
A I S W A C R P P L H P L G Q R A
3651 CGAGGCATAATAGGGCTCGGCCGCGTCAAGCAGGCGGATGATGACGGAAT 3700
GCTCCGTATTATCCCGAGCCGGCGCAGTTCGTCCGCCTACTACTGCCTTA
S A Y Y P E A A D L L R I I V S
3701 AGAGCGCGTCCGAAGGCACCGGACCCTCAACCGTCGCCCCCGCCTCGGCC 3750
TCTCGCGCAGGCTTCCGTGGCCTGGGAGTTGGCAGCGGGGGCGGAGCCGG
Y L A D S P V P G E V T A G A E A
3751 AGCCAGTCGGCAGGCAGATAGCAGCGCCCGATGGCGGCATCGTCGATCAC 3800
TCGGTCAGCCGTCCGTCTATCGTCGCGGGCTACCGCCGTAGCAGCTAGTG
L W D A P L Y C R G I A A D D I V
3801 GTCGCGAGCGATGTTCGTCAGCTGGAACGCAAGGCCCAGATCGCAGGCGC 3850
CAGCGCTCGCTACAAGCAGTCGACCTTGCGTTCCGGGTCTAGCGTCCGCG
D R A I N T L Q F A L G L D C A
3851 GATCCAGCACCGCATCGTCCTGCACGCCCATCACCCGCGCCATCATCACG 3900
CTAGGTCGTGGCGTAGCAGGACGTGCGGGTAGTGGGCGCGGTAGTAGTGC
R D L V A D D Q V G M V R A M M V
FIG. 7G
3901 CCCACGACCCCCGCGACGTGGTAGGAATATTCCAGCACGTCATCCAGGCT 3950
GGGTGCTGGGGGCGCTGCACCATCCTTATAAGGTCGTGCAGTAGGTCCGA
G V V G A V H Y S Y E L V D D L S
3951 GCGGTATTCGCGATCCGCGACATCCATCGCGAAACCCTCGATCAGGTCCA 4000
CGCCATAAGCGCTAGGCGCTGTAGGTAGCGCTTTGGGAGCTAGTCCAGGT
R Y E R D A V D M A F G E I L D
4001 TCGGCCAAAGGTCCGGGAAATCATGCCGCCGGGCGACCTGGCGCAGCGCC 4050
AGCCGGTTTCCAGGCCCTTTAGTACGGCGGCCCGCTGGACCGCGTCGCGG
M P W L D P F D H R R A V Q R L A
4051 GCGAAGGGCGGCGACATCGGGCCGTCCTCGTGCAGCGCGGCCAGCGTGTC 4100
CGCTTCCCGCCGCTGTAGCCCGGCAGGAGCACGTCGCGCCGGTCGCACAG
A F P P S M P G D E H L A A L T D
4101 GGCGCGCAGCGCCCCCAGCCGCGCCTGTGGGTCGCCGCCCGCCTCGGGGG 4150
CCGCGCGTCGCGGGGGTCGGCGCGGACACCCAGCGGCGGGCGGAGCCCCC
A R L A G L R A Q P D G G A E P
4151 CAGAACCCATCACCTGCCCGTCGATCACGTCATCCGCATGCCTGCACCAG 4200
GTCTTGGGTAGTGGACGGGCAGCTAGTGCAGTAGGCGTACGGACGTGGTC
A S G M V Q G D I V D D A H R C W


4201 GCATAGAGCATGACCGTATCCTCGCGGATGCCGGGCGGCATCAGCTTGGC
CGTATCTCGTACTGGCATAGGAGCGCCTACGGCCCGCCGTAGTCGAACCG
4250
A Y L M V T D E R I G P P M L K A
4251 CGCCTGCGCGAAGCTTTGCGAACCCTGCGCGATGGCCGCTTCGGAAGTCG
GCGGACGCGCTTCGAAACGCTTGGGACGCGCTACCGGCGAAGCCTTCAGC
4300
A Q A F S Q S G Q A I A A E S T
4301 CCGTCAGATCGGTCATGCGACGGCCAGGTCCGACAGCATGACCTGCGCCG
GGCAGTCTAGCCAGTACGCTGCCGGTCCAGGCTGTCGTACTGGACGCGGC
4350
* A V A L D S L M V Q A
A T L D T M
<-- crtB
4351 TGGCCTTGGCGCTGCCAACGACACCCGGGATGCCCGCACCCGGATGCGTG
ACCGGAACCGCGACGGTTGCTGTGGGCCCTACGGGCGTGGGCCTACGCAC
4400
T A K A S G V V G P I G A G P H T
4401 CCCGCCCCCACGATGTAGAAGTTCGGGATCGCGCGGTCGCGGTTATGCGG
GGGCGGGGGTGCTACATCTTCAAGCCCTAGCGCGCCAGCGCCAATACGCC
4450
G A G V I Y F N P I A R D R N H P
4451 GCGGAACCAGGCGGATTGCGTCAGGATCGGCTCGACCGAGAAGGCGCTGC
CGCCTTGGTCCGCCTAACGCAGTCCTAGCCGAGCTGGCTCTTCCGCGACG
4500
R F W A S Q T L I P E V S F A S
FIG. 7H
4501 CGTGATGGGCCGACAGTTCGGTGCTGAAATCGGCGGGGCTGAAGATGCGG
GCACTACCCGGCTGTCAAGCCACGACTTTAGCCGCCCCGACTTCTACGCC
4550
G H H A S L E T S F D A P S F I R
4551 CTGACGGTCAGGTGCTTGCGCAGGTCGGGGATGGCGCGGCGCTCCAGTTC
GACTGCCAGTCCACGAACGCGTCCAGCCCCTACCGCGCCGCGAGGTCAAG
4600
S V T L H K R L D P I A R R E L E
4601 CTCGAAGATGCGCTCGGCATAGCCCGGGGCCTCGGCTTCCCAATCGACAT
GAGCTTCTACGCGAGCCGTATCGGGCCCCGGAGCCGAAGGGTTAGCTGTA
4650
E F I R E A Y G P A E A E W D V
4651 CGGCGCGGCCCAGATGCGGAACGGGCGCAAGGACGTAATGCGTGGACATC
GCCGCGCCGGGTCTACGCCTTGCCCGCGTTCCTGCATTACGCACCTGTAG
4700
D A R G L H P V P A L V Y H T S M
4701 CCCTCGGGGGCCAGGCTGGGATCGGTCACGCAGGGCGAATGCAGATACAT
GGGAGCCCCCGGTCCGACCCTAGCCAGTGCGTCCCGCTTACGTCTATGTA
4750
G E P A L S P D T V C P S H L Y M
4751 CGAGAAATCGTCCGGCAGGCGTGGCCCGTTGAAGATCTCGTTCACCAGCC
GCTCTTTAGCAGGCCGTCCGCACCGGGCAACTTCTAGAGCAAGTGGTCGG
4800
S F D D P L R P G N F I E N V L

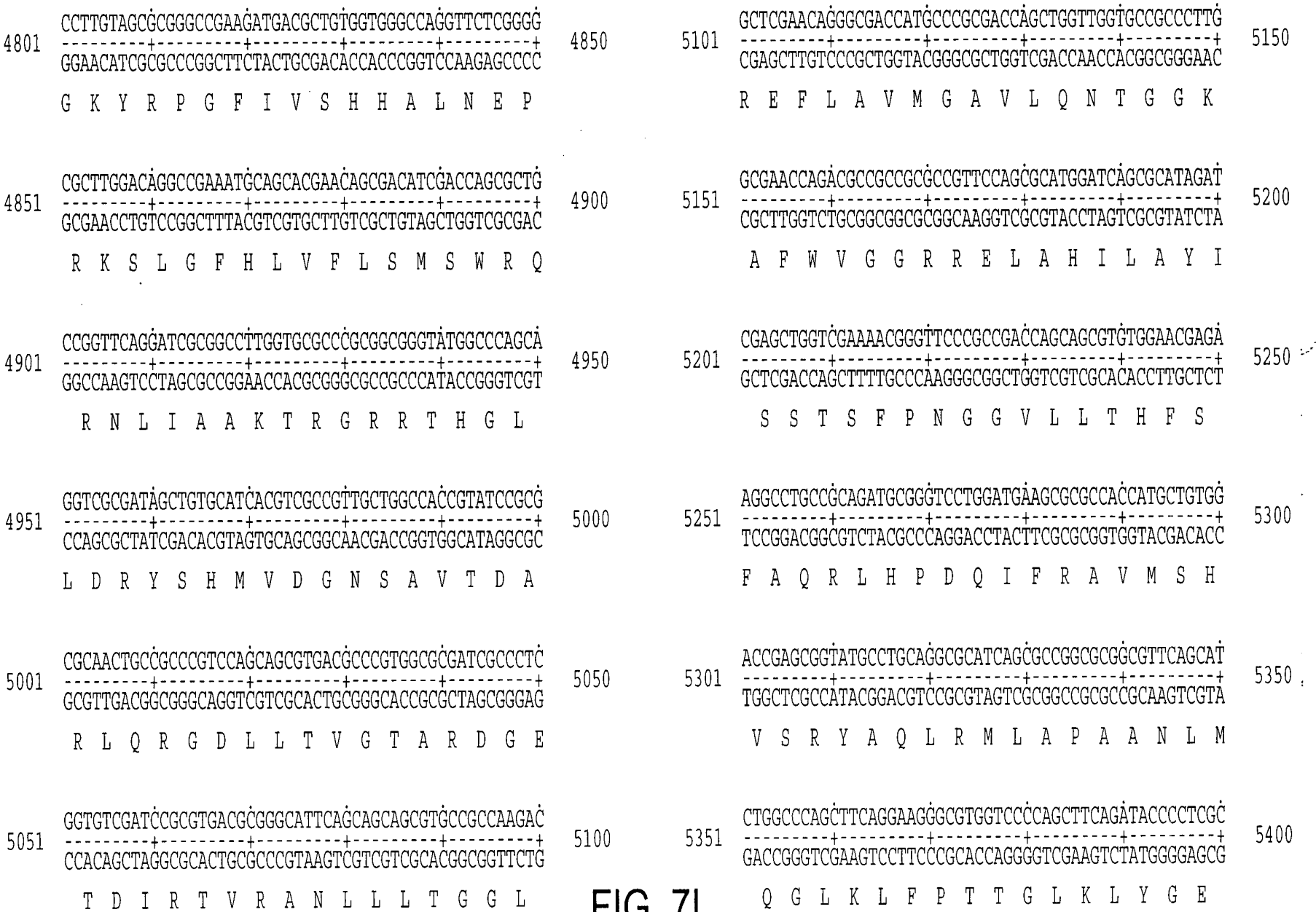

4801 CCTTGTAGCGCGGGCCGAAGATGACGCTGTGGTGGGCCAGGTTCTCGGGG
GGAACATCGCGCCCGGCTTCTACTGCGACACCACCCGGTCCAAGAGCCCC
G K Y R P G F I V S H H A L N E P
4850
4851 CGCTTGGACAGGCCGAAATGCAGCACGAACAGCGACATCGACCAGCGCTG
GCGAACCTGTCCGGCTTTACGTCGTGCTTGTCGCTGTAGCTGGTCGCGAC
R K S L G F H L V F L S M S W R Q
4900
4901 CCGGTTCAGGATCGCGGCCTTGGTGCGCCCGCGGCGGGTATGGCCCAGCA
GGCCAAGTCCTAGCGCCGGAACCACGCGGGCGCCGCCCATACCGGGTCGT
R N L I A A K T R G R R T H G L
4950
4951 GGTCGCGATAGCTGTGCATCACGTCGCCGTTGCTGGCCACCGTATCCGCG
CCAGCGCTATCGACACGTAGTGCAGCGGCAACGACCGGTGGCATAGGCGC
L D R Y S H M V D G N S A V T D A
5000
5001 CGCAACTGCCGCCCGTCCAGCAGCGTGACGCCCGTGGCGCGATCGCCCTC
GCGTTGACGGCGGGCAGGTCGTCGCACTGCGGGCACCGCGCTAGCGGGAG
R L Q R G D L L T V G T A R D G E
5050
5051 GGTGTCGATCCGCGTGACGCGGGCATTCAGCAGCAGCGTGCCGCCAAGAC
CCACAGCTAGGCGCACTGCGCCCGTAAGTCGTCGTCGCACGGCGGTTCTG
T D I R T V R A N L L L T G G L
5100
5101 GCTCGAACAGGGCGACCATGCCCGCGACCAGCTGGTTGGTGCCGCCCTTG
CGAGCTTGTCCCGCTGGTACGGGCGCTGGTCGACCAACCACGGCGGGAAC
R E F L A V M G A V L Q N T G G K
5150
5151 GCGAACCAGACGCCGCCGCGCCGTTCCAGCGCATGGATCAGCGCATAGAT
CGCTTGGTCTGCGGCGGCGCGGCAAGGTCGCGTACCTAGTCGCGTATCTA
A F W V G G R R E L A H I L A Y I
5200
5201 CGAGCTGGTCGAAAACGGGTTCCCGCCGACCAGCAGCGTGTGGAACGAGA
GCTCGACCAGCTTTTGCCCAAGGGCGGCTGGTCGTCGCACACCTTGCTCT
S S T S F P N G G V L L T H F S
5250
5251 AGGCCTGCCGCAGATGCGGGTCCTGGATGAAGCGCGCCACCATGCTGTGG
TCCGGACGGCGTCTACGCCCAGGACCTACTTCGCGCGGTGGTACGACACC
F A Q R L H P D Q I F R A V M S H
5300
5301 ACCGAGCGGTATGCCTGCAGGCGCATCAGCGCCGGCGCGGCGTTCAGCAT
TGGCTCGCCATACGGACGTCCGCGTAGTCGCGGCCGCGCCGCAAGTCGTA
V S R Y A Q L R M L A P A A N L M
5350
5351 CTGGCCCAGCTTCAGGAAGGGCGTGGTCCCCAGCTTCAGATACCCCTCGC
GACCGGGTCGAAGTCCTTCCCGCACCAGGGGTCGAAGTCTATGGGGAGCG
Q G L K L F P T T G L K L Y G E
5400

FIG. 7I

```
5401 GATAGACCTCCTCGGCGTAATCGTGGAAGCGGCGATAGCCATCGACATCG 5450
     CTATCTGGAGGAGCCGCATTAGCACCTTCGCCGCTATCGGTAGCTGTAGC
       R  Y  V  E  E  A  Y  D  H  F  R  R  Y  G  D  V  D

5451 GCGGGATTGAAGGAGGCGACCTGGCGGATCAGCTCGTCGTCGTCGTTCAC 5500
     CGCCCTAACTTCCTCCGCTGGACCGCCTAGTCGAGCAGCAGCAGCAAGTG
       A  P  N  F  S  A  V  Q  R  I  L  E  D  D  D  N  V

5501 GTATTCGAAGCTGCGGCCGTCCGCCCATGTCAGCCGGTAGAAGGGCGAGA 5550
     CATAAGCTTCGACGCCGGCAGGCGGGTACAGTCGGCCATCTTCCCGCTCT
       Y  E  F  S  R  G  D  A  W  T  L  R  Y  F  P  S

5551 CCGGCAGCAGCGTCACGTCACGCTCCATCGGTTGGCCGCTGAGGGCCCAC 5600
     GGCCGTCGTCGCAGTGCAGTGCGAGGTAGCCAACCGGCGACTCCCGGGTG
       V  P  L  L  T  V  D  R  E  M  P  Q  G  S  L  A  W

5601 AGCTCTCGCAGGCTGTCGGGGTCGGTCACGACCGTCGGGCCTGCATCGAA 5650
     TCGAGAGCGTCCGACAGCCCCAGCCAGTGCTGGCAGCCCGGACGTAGCTT
       L  E  R  L  S  D  P  D  T  V  V  T  P  G  A  D  F

5651 GACGTGGCCCTGATCGTTCCAGACATAGGCGCGGCCGCCGGGCTTGTCGC 5700
     CTGCACCGGGACTAGCAAGGTCTGTATCCGCGCCGGCGGCCCGAACAGCG
       V  H  G  Q  D  N  W  V  Y  A  R  G  G  P  K  D

5701 GGGCCTCGACGATGGTGGTCGCGATGCCGGCCGATTGCAGGCGGATGGCA 5750
     CCCGGAGCTGCTACCACCAGCGCTACGGCCGGCTAACGTCCGCCTACCGT
       R  A  E  V  I  T  T  A  I  G  A  S  Q  L  R  I  A

5751 AGCGCAAGCCCGCCGAAACCTGCGCCGATGACGATGGCGGAACTCATGCT 5800
     TCGCGTTCGGGCGGCTTTGGACGCGGCTACTGCTACCGCCTTGAGTACGA
       L  A  L  G  G  F  G  A  G  I  V  I  A  S  S  M  <-- crtI
                                                  *  A

5801 CTCTCCTGCAGCAGGGGGCGTTCGGGCAGGCAGCGCACGGCCTGCGACAG 5850
     GAGAGGACGTCGTCCCCCGCAAGCCCGTCCGTCGCGTGCCGGACGCTGTC
       R  E  Q  L  L  P  R  E  P  L  C  R  V  A  Q  S  L

5851 CGGAATGGGCGGGCGTCCGGTGACGATGCGAAGCCGGTCGGCCAATGTCA 5900
     GCCTTACCCGCCCGCAGGCCACTGCTACGCTTCGGCCAGCCGGTTACAGT
       P  I  P  P  R  G  T  V  I  R  L  R  D  A  L  T

5901 GGCGCCCGGCATAGAAGCGCTCGATCAGCGGCTGCGGCAGGCGGTAGAAC 5950
     CCGCGGGCCGTATCTTCGCGAGCTAGTCGCCGACGCCGTCCGCCATCTTG
       L  R  G  A  Y  F  R  E  I  L  P  Q  P  L  R  Y  F

5951 CGCTGCAGCAGGCGATAGCGACGGTCGGGCGGGCAGCCGCGGAACAGCAT 6000
     GCGACGTCGTCCGCTATCGCTGCCAGCCCGCCCGTCGGCGCCTTGTCGTA
       R  Q  L  L  R  Y  R  R  D  P  P  C  G  R  F  L  M
```

FIG. 7J

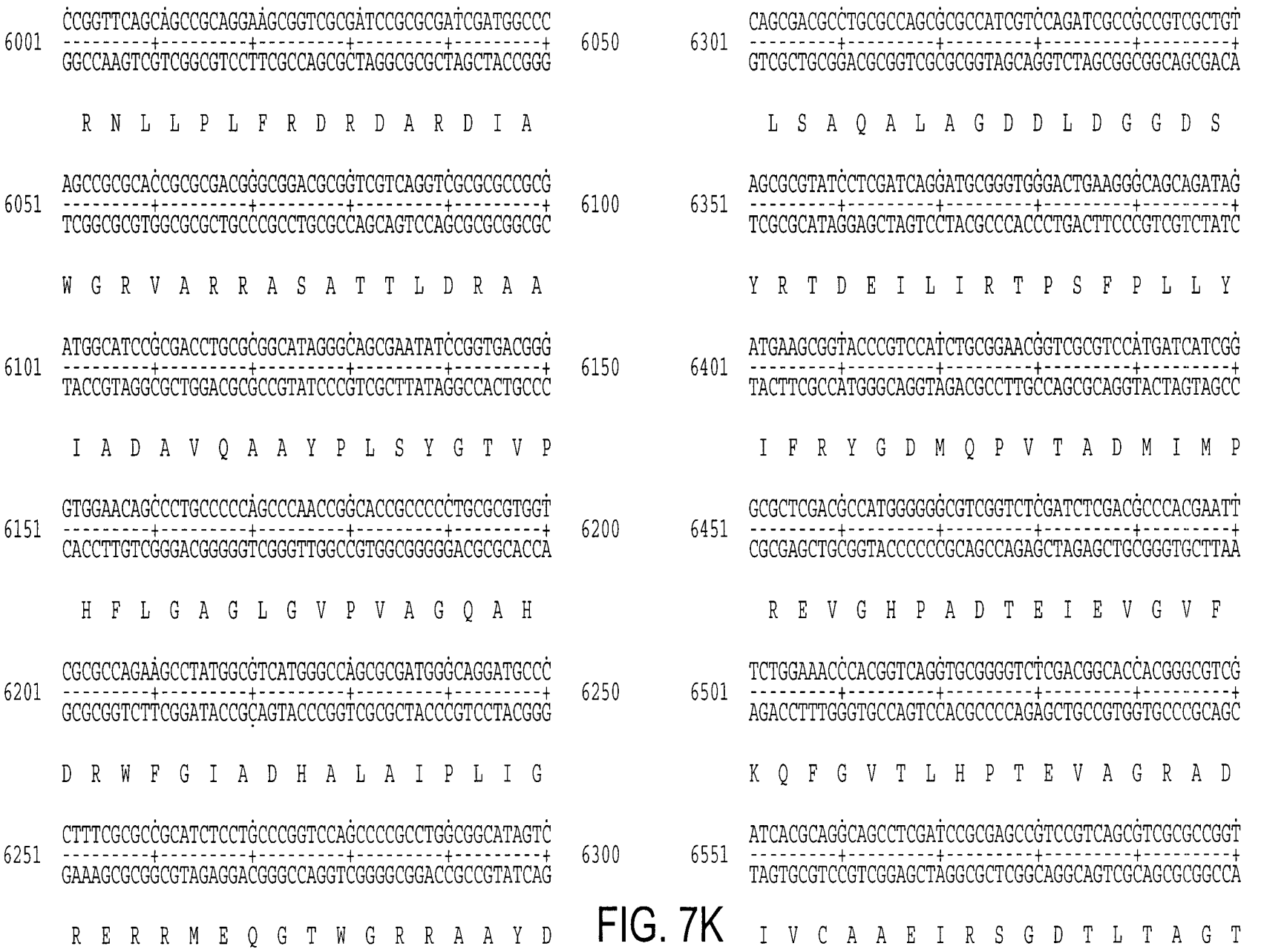

6001 CCGGTTCAGCAGCCGCAGGAAGCGGTCGCGATCCGCGCGATCGATGGCCC 6050
GGCCAAGTCGTCGGCGTCCTTCGCCAGCGCTAGGCGCGCTAGCTACCGGG
R N L L P L F R D R D A R D I A
6051 AGCCGCGCACCGCGCGACGGGCGGACGCGGTCGTCAGGTCGCGCGCCGCG 6100
TCGGCGCGTGGCGCGCTGCCCGCCTGCGCCAGCAGTCCAGCGCGCGGCGC
W G R V A R R A S A T T L D R A A
6101 ATGGCATCCGCGACCTGCGCGGCATAGGGCAGCGAATATCCGGTGACGGG 6150
TACCGTAGGCGCTGGACGCGCCGTATCCCGTCGCTTATAGGCCACTGCCC
I A D A V Q A A Y P L S Y G T V P
6151 GTGGAACAGCCCTGCCCCCAGCCCAACCGGCACCGCCCCCTGCGCGTGGT 6200
CACCTTGTCGGGACGGGGGTCGGGTTGGCCGTGGCGGGGGACGCGCACCA
H F L G A G L G V P V A G Q A H
6201 CGCGCCAGAAGCCTATGGCGTCATGGGCCAGCGCGATGGGCAGGATGCCC 6250
GCGCGGTCTTCGGATACCGCAGTACCCGGTCGCGCTACCCGTCCTACGGG
D R W F G I A D H A L A I P L I G
6251 CTTTCGCGCCGCATCTCCTGCCCGGTCCAGCCCCGCCTGGCGGCATAGTC 6300
GAAAGCGCGGCGTAGAGGACGGGCCAGGTCGGGGCGGACCGCCGTATCAG
R E R R M E Q G T W G R R A A Y D
6301 CAGCGACGCCTGCGCCAGCGCGCCATCGTCCAGATCGCCGCCGTCGCTGT 6350
GTCGCTGCGGACGCGGTCGCGCGGTAGCAGGTCTAGCGGCGGCAGCGACA
L S A Q A L A G D D L D G G D S
6351 AGCGCGTATCCTCGATCAGGATGCGGGTGGGACTGAAGGGCAGCAGATAG 6400
TCGCGCATAGGAGCTAGTCCTACGCCCACCCTGACTTCCCGTCGTCTATC
Y R T D E I L I R T P S F P L L Y
6401 ATGAAGCGGTACCCGTCCATCTGCGGAACGGTCGCGTCCATGATCATCGG 6450
TACTTCGCCATGGGCAGGTAGACGCCTTGCCAGCGCAGGTACTAGTAGCC
I F R Y G D M Q P V T A D M I M P
6451 GCGCTCGACGCCATGGGGGGCGTCGGTCTCGATCTCGACGCCCACGAATT 6500
CGCGAGCTGCGGTACCCCCCGCAGCCAGAGCTAGAGCTGCGGGTGCTTAA
R E V G H P A D T E I E V G V F
6501 TCTGGAAACCCACGGTCAGGTGCGGGGTCTCGACGGCACCACGGGCGTCG 6550
AGACCTTTGGGTGCCAGTCCACGCCCCAGAGCTGCCGTGGTGCCCGCAGC
K Q F G V T L H P T E V A G R A D
6551 ATCACGCAGGCAGCCTCGATCCGCGAGCCGTCCGTCAGCGTCGCGCCGGT 6600
TAGTGCGTCCGTCGGAGCTAGGCGCTCGGCAGGCAGTCGCAGCGCGGCCA
I V C A A E I R S G D T L T A G T

FIG. 7K

6601
ATCGTCCAGCGTCGCGACATGCGTATTCCACCGCAGATCGACACCCTGCA
TAGCAGGTCGCAGCGCTGTACGCATAAGGTGGCGTCTAGCTGTGGGACGT
6650
D D L T A V H T N W R L D V G Q
6651
GCAGCCCGATCAGCGCGCCCGCCTCGATCGAGCCATAGCCTGTCGTCAGG
CGTCGGGCTAGTCGCGCGGGCGGAGCTAGCTCGGTATCGGACAGCAGTCC
6700
L L G I L A G A E I S G Y G T T L
6701
CGGCGCGAATGGTCGGGAAACGCGACCTCCTGATCCGTCCATTCGCCGCG
GCCGCGCTTACCAGCCCTTTGCGCTGGAGGACTAGGCAGGTAAGCGGCGC
6750
R R S H D P F A V E Q D T W E G R
6751
ACGAATGGGCGACAGGCGCGCCAGCCATTCGGGCGAAAGATCCGTGTCGT
TGCTTACCCGCTGTCCGCGGTCGGTAAGCCCGCTTTCTAGGCACAGCA
6800
R I P S L R A L W E P S L D T D
6801
GGCAGGACCAGGTGTGCTGGTCCGAGGGGCCGGACCGCGCGTCGAGCATC
CCGTCCTGGTCCACACGACCAGGCTCCCCGGCCTGGCGCGCAGCTCGTAG
6850
H C S W T H Q D S P G S R A D L M
6851
ACGATGCGCGCATCCGGTCTGCGGTCGCGAACGGCAAGCGCGATCAGCGC
TGCTACGCGCGTAGGCCAGACGCCAGCGCTTGCCGTTCGCGCTAGTCGCG
6900
V I R A D P R R D R V A L A I L A
6901
ACCGGACAGGCCCCGCCGCCCGCGATCAGCAGATCATGGCTCATGTATTGCG
TGGCCTGTCCGGGGCCGGGCGCTAGTCGTCTAGTACCGAGTACATAACGC
6950
* T N R
G S L G A G A I L L D H S M
<-- crtY
6951
ATCCGCCCCCTTCGCGGTCCTTCAGCAGCGCGCCCGAGGTTTCAGCTCTG
TAGGCGGGGAAGCGCCAGGAAGTCGTCGCGCGGGCTCGCAAAGTCGAGAC
7000
D A G E R D K L L A G S R K L E
7001
CCTTGAGGCTGTCGACCGAGGGCGCCCAGATGAAACCGAGCTGACGCAG
GGAACTCCGACAGCTGGCTCCCGCGGGTCTACTTTGGCTTCGACTGCGTC
7050
A K L S D V S P A W I F G F S V C
7051
TTCTCGCGGCCATGGACCGCGTGATGCATCCTGTGTGCCTGGTAGACGCG
AAGAGCGCCGGTACCTGGCGCACTACGTAGGACACACGGACCATCTGCGC
7100
N E R G H V A H H M R H A Q Y V R
7101
ACGAAGATAGCCGCGCTTGGGGACATAGCGGAACGGCCAGCGCCCATGCA
TGCTTCTATCGGCGCGAACCCCTGTATCGCCTTGCCGGTCGCGGGTACGT
7150
R L Y G R K P V Y R F P W R G H
7151
CCAAGCCGTCATGCAGGAAATAGTAGATCAGCCCGTAGCAGGTGACCCCC
GGTTCGGCAGTACGTCCTTTATCATCTAGTCGGGCATCGTCCACTGGGGG
7200
V L G D H L F Y Y I L G Y C T V G

FIG. 7L

```
7201  ACCGCCAGCCACCAGGCCAGATCCGACCCCATCGCCGCGATCGCGAACAG      7250
      ----------+---------+---------+---------+---------+----------
      TGGCGGTCGGTGGTCCGGTCTAGGCTGGGGTAGCGGCGCTAGCGCTTGTC
        V  A  L  W  W  A  L  D  S  G  G  M  A  G  I  A  F  L

7251  CACGATCGAGATTACCGCGAAGATGACGCCATAGAGGTCGTTCTTCTCGA      7300
      ----------+---------+---------+---------+---------+----------
      GTGCTAGCTCTAATGGCGCTTCTACTGCGGTATCTCCAGCAAGAAGAGCT
        V  I  S  I  V  A  F  I  V  G  Y  L  D  N  K  E

7301  GCGCGTGGTCGTGATCCTCGTCGTGGTGCGATTTATGCCAGCCCCAGCCC      7350
      ----------+---------+---------+---------+---------+----------
      CGCGCACCAGCACTAGGAGCAGCACCACGCTAAATACGGTCGGGGTCGGG
        L  A  H  D  H  D  E  D  H  H  S  K  H  W  G  W  G

7351  AGGGGGCCATGCATGATCCACCGATGGACGGAGTAGGCCGTCAGCTCCATT      7400
      ----------+---------+---------+---------+---------+----------
      TCCCCCCGGTACGTACTAGGTGGCTACCTGCCTCATCCGGCAGTCGAGGTA
        L  P  G  H  M  I  W  R  H  V  S  Y  A  T  L  E  M

7401  CGCGGCGACGGTCAGGATGACGGTCAGGATTGCGGCCCCAAGTGCTCATGC      7450
      ----------+---------+---------+---------+---------+----------
      GCGCCGCTGCCAGTCCTACTGCCAGTCCTAACGCCGGGTTCACGAGTACG
        A  A  V  T  L  I  I  V  T  L  I  A  A  W  T  S  M
                                          <-- crtZ

7451  CGGCCCCTTGCTTGATATGACAGGGAACAGGCTACGCTGCCGCGCGGTGC      7500
      ----------+---------+---------+---------+---------+----------
      GCCGGGGAACGAACTATACTGTCCCTTGTCCGATGCGACGGCGCGCCACG
```

```
7550  ATGACCAGCCCATCGGGTGCGACCAAGGGCATCGCGTGACATCTGCGT      7601
      ----------+---------+---------+---------+---------+----------
      TACTGGTCGGGTAGCCCCACGCTGGTTTCCCGTAGCGCACTGTAGACGCA

7600  TCAGGGCTCATAGGCGGATCATCCGTGACATTCGCCGCCGAACGCGGCAG      7651
      ----------+---------+---------+---------+---------+----------
      AGTCCCGAGTATCCGCCTAGTAGGCACTGTAAGCGGCGGCTTGCGCCGTC

7650  GCGCATCACGCGGTTCCGTCGCTGGAAATATTAATGTTTTCCCGAAGATGG      7701
      ----------+---------+---------+---------+---------+----------
      CGCGTAGTGCGCAAGGCAGCGACCTTTATAATTACAAAAGGGCTTCTACC

7700  TCGGGGCGAGAGGATTCGAACCTCCGACCTACGGTACCCAAAACCGTCGC      7751
      ----------+---------+---------+---------+---------+----------
      AGCCCCGCTCTCCTAAGCTTGGAGGCTGGATGCCATGGGTTTTGGCAGCG

7750  GCTACCAGGCTGGCGCTACGCCCCGACTGCGGAAGGCTTTAGCCGATTGTT      7801
      ----------+---------+---------+---------+---------+----------
      CGATGGTCCGACGCGATGCGGGGCTGACGCCTTCCGAAATCGGCTAACAA

7800  CCGGCAAGGGAAAGACCTAGTCGCAGGCCAGGACCGCATTGTCGCCCATG      7851
      ----------+---------+---------+---------+---------+----------
      GGCCGTTCCCTTTCTGGATCAGCGTCCGGTCCTGGCGTAACAGCGGGTAC

      * D  C  A  L  V  A  N  D  G  M
```

FIG. 7M

```
7801  CCCGGATGCGCCATCGGCTGACCGGGCTTCAGGCCAAGGCGATCCGCCTC  7850
      ---------+---------+---------+---------+---------+
      GGGCCTACGCGGTAGCCGACTGGCCCGAAGTCCGGTTCCGCTAGGCGGAG

          G  P  H  A  M  P  Q  G  G  P  K  L  G  L  R  D  A  E

7851  TCCGCCCGCGATTTCGAGGACGAACAGCCGGTCGGGGTCCGGATCGCCGA  7900
      ---------+---------+---------+---------+---------+
      AGGCGGGCGCTAAAGCTCCTGCTTGTCGGCCAGCCCCAGGCCTAGCGGCT

          G  G  A  I  E  L  V  F  L  R  D  P  D  P  D  G

7901  CCGCCGCGCCCGGAATGGGCGTCTCGTCCAGCGGGCGCGCATTGCCGTGG  7950
      ---------+---------+---------+---------+---------+
      GGCGGCGCGGGCCTTACCCGCAGAGCAGGTCGCCCGCGCGTAACGCCACC

          V  A  A  G  P  I  P  T  E  D  L  P  R  A  N  R  H

7951  ATGTGGCGGATGACGCCGGTTTCATCCGCAAAGACCATGTCCAGCGGGAT  8000
      ---------+---------+---------+---------+---------+
      TACACCGCCTACTGCGGCCAAAGTAGGCGTTTCTGGTACAGGTCGCCCTA

          I  H  R  I  V  G  T  E  D  A  F  V  M  D  L  P  I

8001  CAGTGTGTTGCGCATCCAGAAGGACACCGGCTGGGGCGATTCGTAGATGA  8050
      ---------+---------+---------+---------+---------+
      GTCACACAACGCGTAGGTCTTCCTGTGGCCGACCCCGCTAAGCATCTACT

          L  T  N  R  M  W  F  S  V  P  Q  P  S  E  Y  I

8051  ACAGCATTCCGGTGCCCGCAGGCAGCTCCTTGCGGAACATCAGGCCTGC  8100
      ---------+---------+---------+---------+---------+
      TGTCGTAAGGCCACGGGCGTCCGTCGAGGAACGCCTTGTAGTCCGGGACG

8101  GCGCGCTCTTCGGGGCTGTCCGCGACCTCGACCCGAAACCCGAGCGTTTC  8150
      ---------+---------+---------+---------+---------+
      CGCGCGAGAAGCCCCGACAGGCGCTGGAGCTGGGCTTTGGGCTCGCAAAG

          A  R  E  E  E  P  S  D  A  V  E  V  R  F  G  L  T  E

8151  CGCACCGGTATCGACGACAAGACTGCCGGGCGGCGCATTCCACCGCCGCG  8200
      ---------+---------+---------+---------+---------+
      GCGTGGCCATAGCTGCTGTTCTGACGGCCCGCGCGTAAGGTGGCGGCGC

          A  G  T  D  V  V  L  S  G  P  A  C  E  V  A  A

8201  CGGCGGCGGGCATCAGGACCGCAAGAAGCGCTGCGGCCTTACTCGGCCAC  8250
      ---------+---------+---------+---------+---------+
      GCCGCCGCCCGTAGTCCTGGCGTTCTTCGCGACGCCGGAATGAGCCGGTG

          A  A  A  P  M  L  V  A  L  L  A  A  A  K  S  P  W

8251  ATGGGCAAGATAGGACTGCTCGGCGCCGAGATCCTGCTGACCCTGCGCAT  8300
      ---------+---------+---------+---------+---------+
      TACCCGTTCTATCCTGACGAGCCGCGGCTCTAGGACGACTGGGACGCGTA

          M  P  L  I  P  S  S  P  A  S  I  R  S  V  R  R  M

8301  CCTCGTTCCGGTCATGCAGCGCCAGGTCCATGCGCGATCTGCGCGnnC  8350
      ---------+---------+---------+---------+---------+
      GGAGCAAGGCCAGTACGTCGCGGTCCAGGGTACGGCGCTAGACGCGCnnG

          R  T  G  T  M  <-- orf-16

8351  ATCAGCCCGCGCGGACCCTCGACGACGCGGAGGCAGATCGCCTCGCCGAT  8400
      ---------+---------+---------+---------+---------+
      TAGTCGGGCGCGCCTGGGAGCTGCTGCGCCTCCGTCTAGCGGAGCGGCTA

          F  L  M  G  T  G  A  P  L  E  K  R  F  M  L  G  Q
```

FIG. 7N

```
8401  CACGAGGTCCGAGAAGCCGGAATGACGGAGCACCTCGATATGGATGAACA  8450
      --------+---------+---------+---------+---------+
      GTGCTCCAGGCTCTTCGGCCTTACTGCCTCGTGGAGCTATACCTACTTGT

8451  CGTCCTCGGGGTGGCCGAAGATGTTGGCGAACCGGGAAAAGGCCCTTGGC  8500
      --------+---------+---------+---------+---------+
      GCAGGAGCCCCACCGGCTTCTACAACCGCTTGGCCCTTTTCCGGGAACCG

8501  CTTGTCGAACCACTTGACGCGGGCCGGACGCAGCGGCAnnCGTCCAGATG  8550
      --------+---------+---------+---------+---------+
      GAACAGCTTGGTGAACTGCGCCCGGCCTGCGTCGCCGTnnGCAGGTCTAC

8551  CTCGATCACCTCGGCATCCAGATCGGCGATnGGGGGGTGnCnGTCGCTTT  8600
      --------+---------+---------+---------+---------+
      GAGCTAGTGGAGCCGTAGGTCTAGCCGCTAnCCCCCCACnGnCAGCGAAA

8601  CnnnCGGTTCGATCGACAGGACCTC  8625
      --------+---------+-----
      GnnnGCCAAGCTAGCTGTCCTGGAG
```

FIG. 7O

```
  1  MTPKQQFPLR DLVEIRLAQI SGQFGVVSAP LGAAMSDAAL SPGKRFRAVL
 51  MLMVAESSGG VCDAMVDAAC AVEMVHAASL IFDDMPCMDD ARTRRGQPAT
101  HVAHGEGRAV LAGIALITEA MRILGEARGA TPDQRARLVA SMSRAMGPVG
151  LCAGQDLDLH APKDAAGIER EQDLKTGVLF VAGLEMLSII KGLDKAETEQ
201  LMAFGRQLGR VFQSYDDLLD VIGDKASTGK DTARDTAAPG PKGGLMAVGQ
251  MGDVAQHYRA SRAQLDELMR TRLFRGGQIA DLLARVLPHD IRRSA
```

FIG. 8

```
  1  MTDLTATSEA AIAQGSQSFA QAAKLMPPGI REDTVMLYAW CRHADDVIDG
 51  QVMGSAPEAG GDPQARLGAL RADTLAALHE DGPMSPPFAA LRQVARRHDF
101  PDLWPMDLIE GFAMDVADRE YRSLDDVLEY SYHVAGVVGV MMARVMGVQD
151  DAVLDRACDL GLAFQLTNIA RDVIDDAAIG RCYLPADWLA EAGATVEGPV
201  PSDALYSVII RLLDAAEPYY ASARQGLPHL PPRCAWSIAA ALRIYRAIGT
251  RIRQGGPEAY RQRISTSKAA KIGLLARGGL DAAASRLRGG EISRDGLWTR
301  PRA
```

FIG. 9

```
  1  MSSAIVIGAG FGGLALAIRL QSAGIATTIV EARDKPGGRA YVWNDQGHVF
 51  DAGPTVVTDP DSLRELWALS GQPMERDVTL LPVSPFYRLT WADGRSFEYV
101  NDDDELIRQV ASFNPADVDG YRRFHDYAEE VYREGYLKLG TTPFLKLGQM
151  LNAAPALMRL QAYRSVHSMV ARFIQDPHLR QAFSFHTLLV GGNPFSTSSI
201  YALIHALERR GGVWFAKGGT NQLVAGMVAL FERLGGTLLL NARVTRIDTE
251  GDRATGVTLL DGRQLRADTV ASNGDVMHSY RDLLGHTRRG RTKAAILNRQ
301  RWSMSLFVLH FGLSKRPENL AHHSVIFGPR YKGLVNEIFN GPRLPDDFSM
351  YLHSPCVTDP SLAPEGMSTH YVLAPVPHLG RADVDWEAEA PGYAERIFEE
401  LERRAIPDLR KHLTVSRIFS PADFSTELSA HHGSAFSVEP ILTQSAWFRP
451  HNRDRAIPNF YIVGAGTHPG AGIPGVVGSA KATAQVMLSD LAVA
```

FIG. 10

```
  1  MSHDLLIAGA GLSGALIALA VRDRRPDARI VMLDARSGPS DQHTWSCHDT
 51  DLSPEWLARL SPIRRGEWTD QEVAFPDHSR RLTTGYGSIE AGALIGLLQG
101  VDLRWNTHVA TLDDTGATLT DGSRIEAACV IDARGAVETP HLTVGFQKFV
151  GVEIETDAPH GVERPMIMDA TVPQMDGYRF IYLLPFSPTR ILIEDTRYSD
201  GGDLDDGALA QASLDYAARR GWTGQEMRRE RGILPIALAH DAIGFWRDHA
251  QGAVPVGLGA GLFHPVTGYS LPYAAQVADA IAARDLTTAS ARRAVRGWAI
301  DRADRDRFLR LLNRMLFRGC PPDRRYRLLQ RFYRLPQPLI ERFYAGRLTL
351  ADRLRIVTGR PPIPLSQAVR CLPERPLLQE RA
```

FIG. 11

```
  1  MSTWAAILTV ILTVAAMELT AYSVHRWIMH GPLGWGWHKS HHDEDHDHAL
 51  EKNDLYGVIF AVISIVLFAI GAMGSDLAWW LAVGVTCYGL IYYFLHDGLV
101  HGRWPFRYVP KRGYLRRVYQ AHRMHHAVHG RENCVSFGFI WAPSVDSLKA
151  ELKRSGALLK DREGADRNT
```

FIG. 12

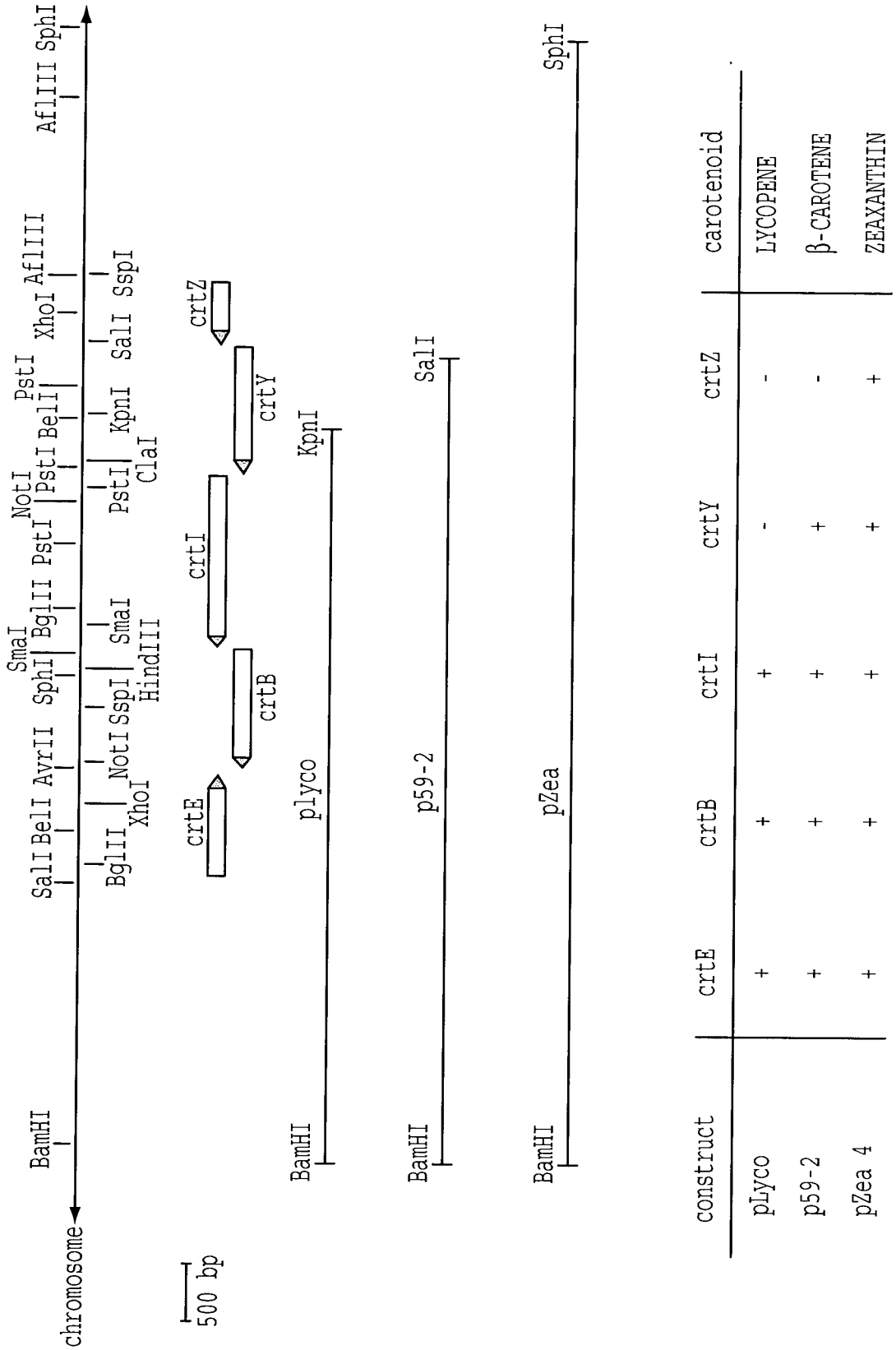

chromosome
BamHI
SalI BclI AvrII NotI XhoI BglII SspI HindIII SphI SmaI BglII SmaI PstI NotI PstI ClaI PstI BclI PstI KpnI SalI SspI XhoI AflIII AflIII SphI
crtE
crtB
crtI
crtY
crtZ
500 bp
plyco
p59-2
pZea
KpnI
SalI
SphI
construct
carotenoid
pLyco
p59-2
pZea 4
LYCOPENE
β-CAROTENE
ZEAXANTHIN

FIG. 13

crtE
#100: 5'tatatactagtaagaggagaaattacatATGACGCCCAAGCAGCAGCAATTC 3'
SpeI
RBS
NdeI
#101: 5'TATATACCCGGGTCAGCCGCGACGGCCTGTGG 3'
SmaI
crtZ
#104: 5'tatatgaattcaagaggagaaattacatATGAGCACTTGGGCCGCAATCC 3'
EcoRI
RBS
NdeI
#105: 5'GTTTCAGCTCTGCCTTGAGGC 3'
crtZ
crtY
MUT1: 5'GCGAAGGGGCGGGATCGCAATACgTGaaaggaggacacgtgATGAGCCATGATCTGCTGATCG 3'
PmlI
crtY
crtI
MUT2: 5'GCCCCCTGCTGCAGGAGAGAGCtTGaaaggaggcaattgagATGAGTTCCGCCATCGTCATCG 3'
MunI
crtI
crtB
MUT3: 5'GGTCATGCTGTCGGACCTGGCCGTCGCtTGaaaggaggatccaatcATGACCGATCTGACGGCGACTTCC 3'
BamHI
MUT5: 5'ATATATctcaattgcctcctttcaaGCTCTCTCCTGCAGCAGGG 3'
MunI
crtY
MUT6: 5'atgattggatcctcctttcaaGCGACGGCCAGGTCCGACAGC 3'
BamHI
crtI
CAR17: 5'CAGAACCCATCACCTGCCCGTC 3'
cat3: 5'CGCGAATTCTCGCCGGCAATAGTTACC 3'
EcoRI
cat4: 5'GTCACATGCATGCATGTTACGAGCTCATAAGCATGTGACGTCTTCAACTAACGGGGCAGG 3'
SphI
SacI
AatII

FIG. 14

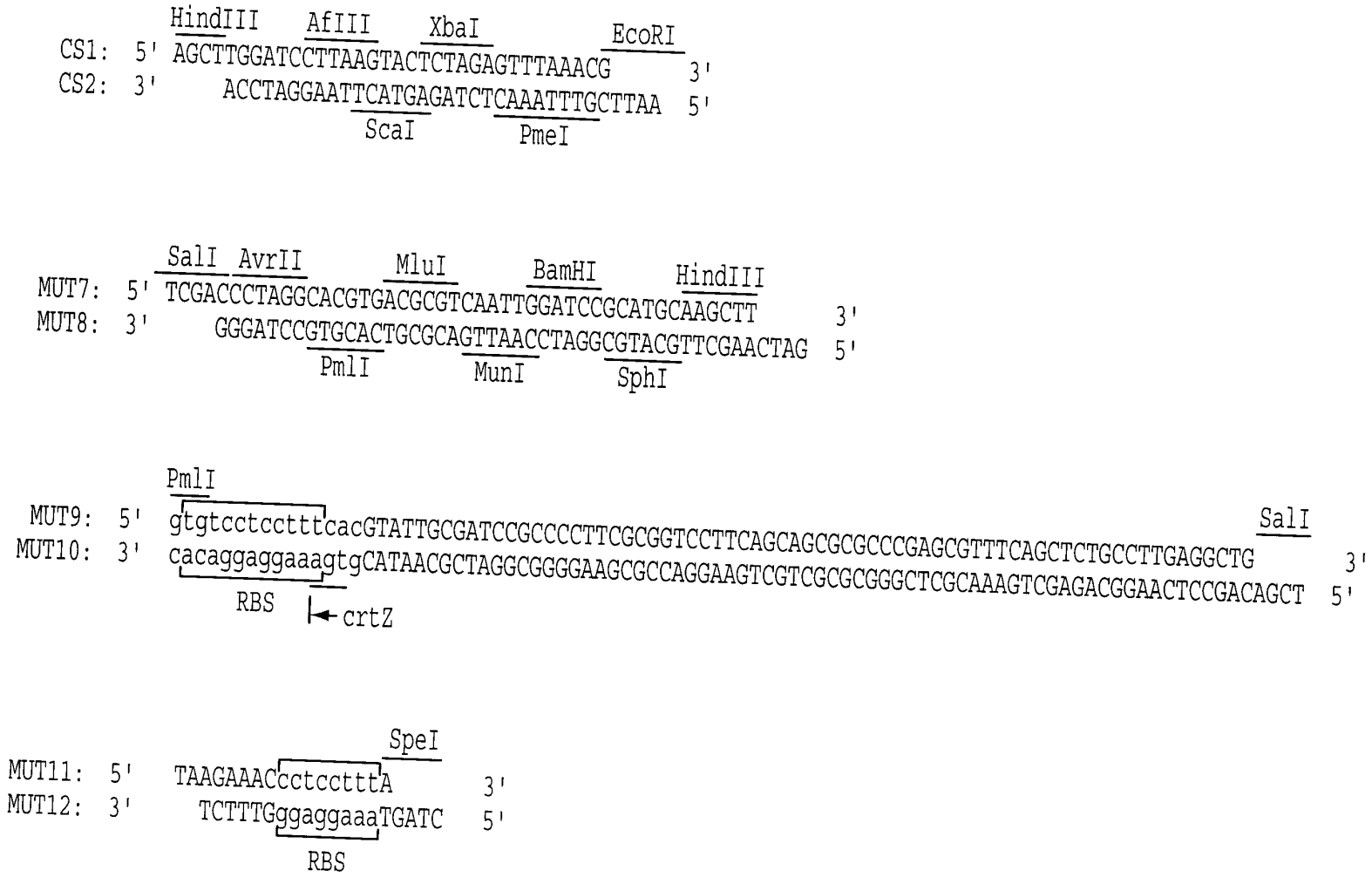

HindIII
AfIII
XbaI
EcoRI
CS1: 5' AGCTTGGATCCTTAAGTACTCTAGAGTTTAAACG 3'
CS2: 3' ACCTAGGAATTCATGAGATCTCAAATTTGCTTAA 5'
ScaI
PmeI
SalI
AvrII
MluI
BamHI
HindIII
MUT7: 5' TCGACCCTAGGCACGTGACGCGTCAATTGGATCCGCATGCAAGCTT 3'
MUT8: 3' GGGATCCGTGCACTGCGCAGTTAACCTAGGCGTACGTTCGAACTAG 5'
PmlI
MunI
SphI
PmlI
SalI
MUT9: 5' gtgtcctcctttcacGTATTGCGATCCGCCCCTTCGCGGTCCTTCAGCAGCGCGCCCGAGCGTTTCAGCTCTGCCTTGAGGCTG 3'
MUT10: 3' cacaggaggaaagtgCATAACGCTAGGCGGGGAAGCGCCAGGAAGTCGTCGCGCGGGCTCGCAAAGTCGAGACGGAACTCCGACAGCT 5'
RBS
crtZ
SpeI
MUT11: 5' TAAGAAACcctcctttA 3'
MUT12: 3' TCTTTGggaggaaaTGATC 5'
RBS

FIG. 15

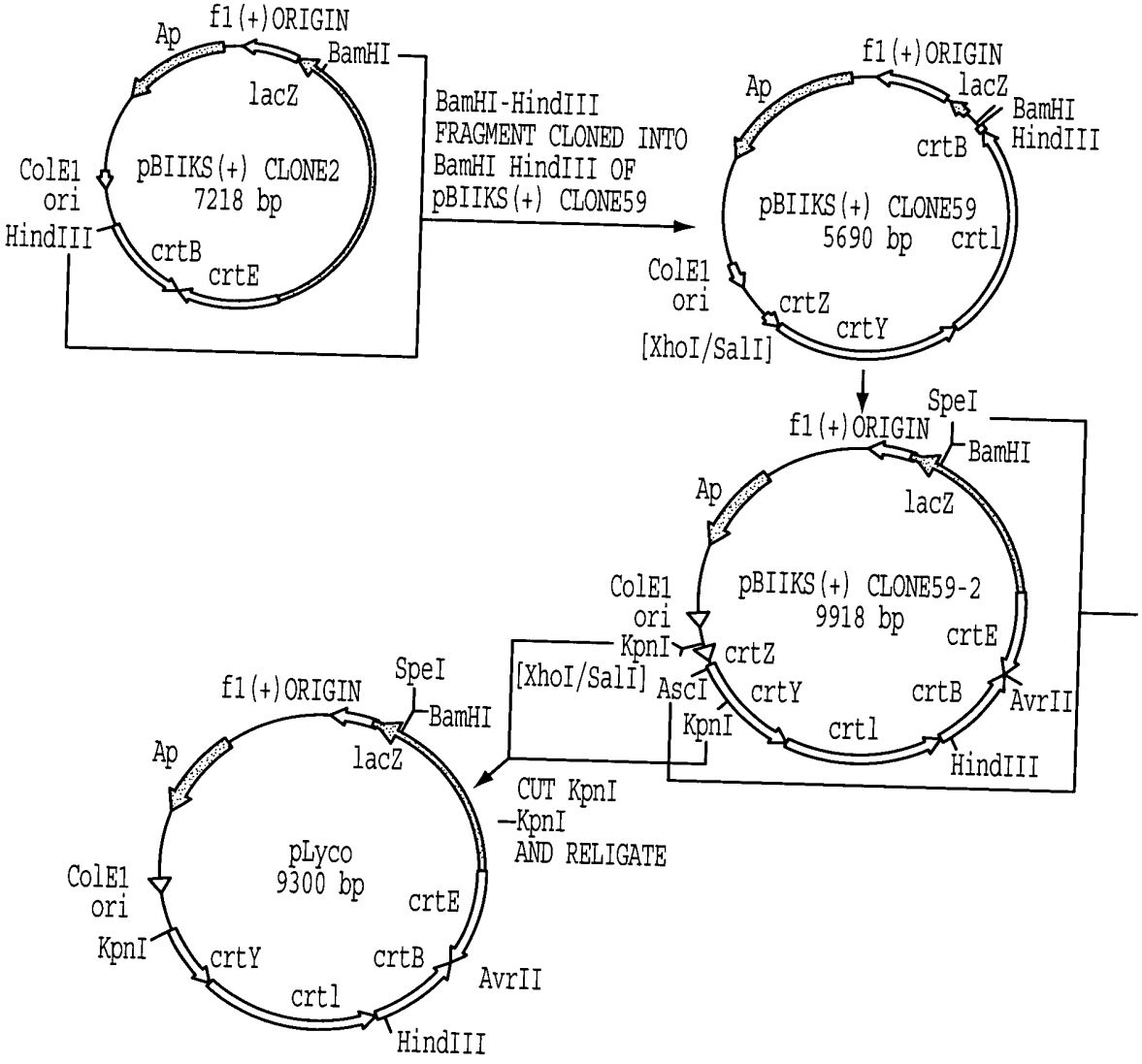

f1(+)ORIGIN
Ap
BamHI
lacZ
ColE1
ori
pBIIKS(+) CLONE2
7218 bp
HindIII
crtB
crtE
BamHI-HindIII
FRAGMENT CLONED INTO
BamHI HindIII OF
pBIIKS(+) CLONE59
f1(+)ORIGIN
Ap
lacZ
BamHI
HindIII
crtB
pBIIKS(+) CLONE59
5690 bp
crtI
ColE1
ori
crtZ
crtY
[XhoI/SalI]
SpeI
f1(+)ORIGIN
BamHI
Ap
lacZ
ColE1
ori
pBIIKS(+) CLONE59-2
9918 bp
KpnI
crtZ
crtE
[XhoI/SalI]
AscI
crtY
crtB
AvrII
KpnI
crtI
HindIII
SpeI
f1(+)ORIGIN
BamHI
Ap
lacZ
CUT KpnI
KpnI
AND RELIGATE
pLyco
9300 bp
ColE1
ori
crtE
KpnI
crtY
crtB
AvrII
crtI
HindIII

FIG. 16A

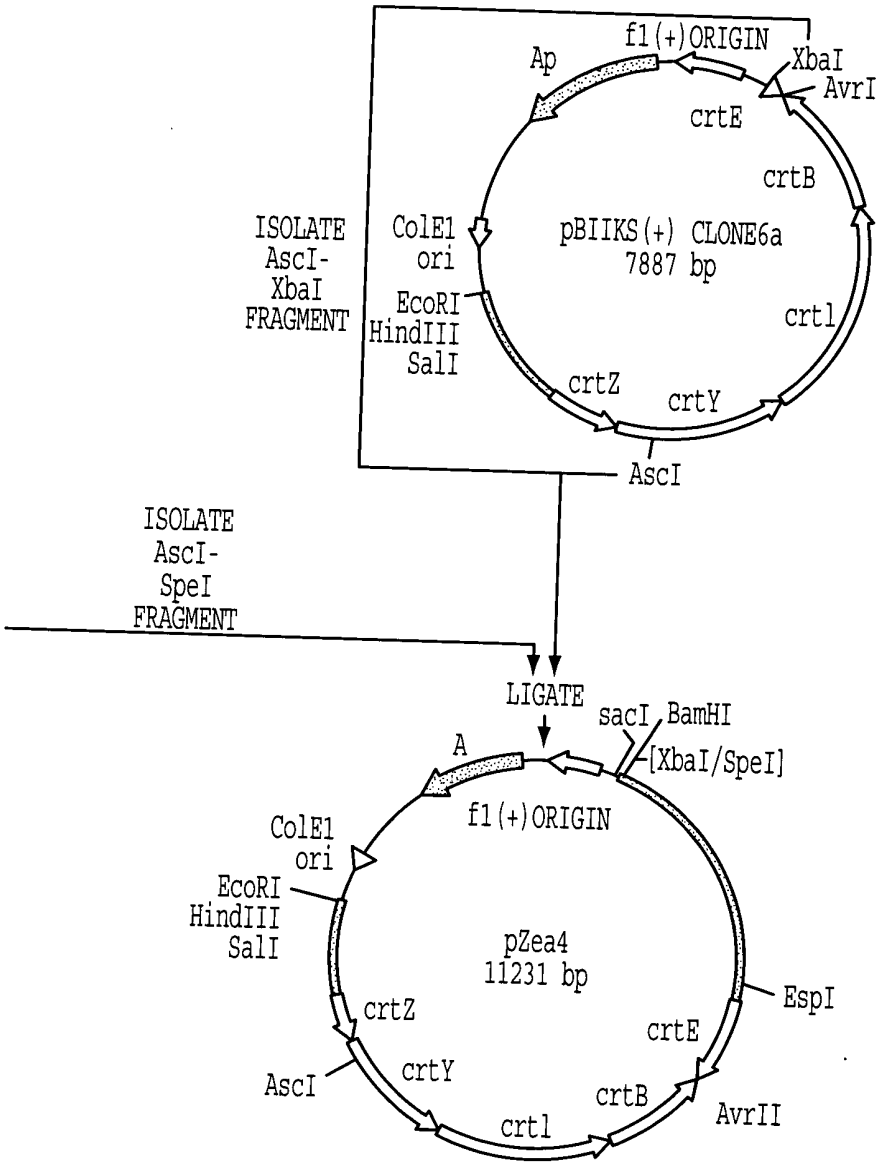

f1(+)ORIGIN
Ap
XbaI
AvrI
crtE
crtB
ISOLATE
AscI-
XbaI
FRAGMENT
ColE1
ori
pBIIKS(+) CLONE6a
7887 bp
EcoRI
HindIII
SalI
crtI
crtZ
crtY
AscI
ISOLATE
AscI-
SpeI
FRAGMENT
LIGATE
sacI
BamHI
[XbaI/SpeI]
A
f1(+)ORIGIN
ColE1
ori
EcoRI
HindIII
SalI
pZea4
11231 bp
EspI
crtZ
crtE
crtY
AscI
crtB
AvrII
crtI

FIG. 16B

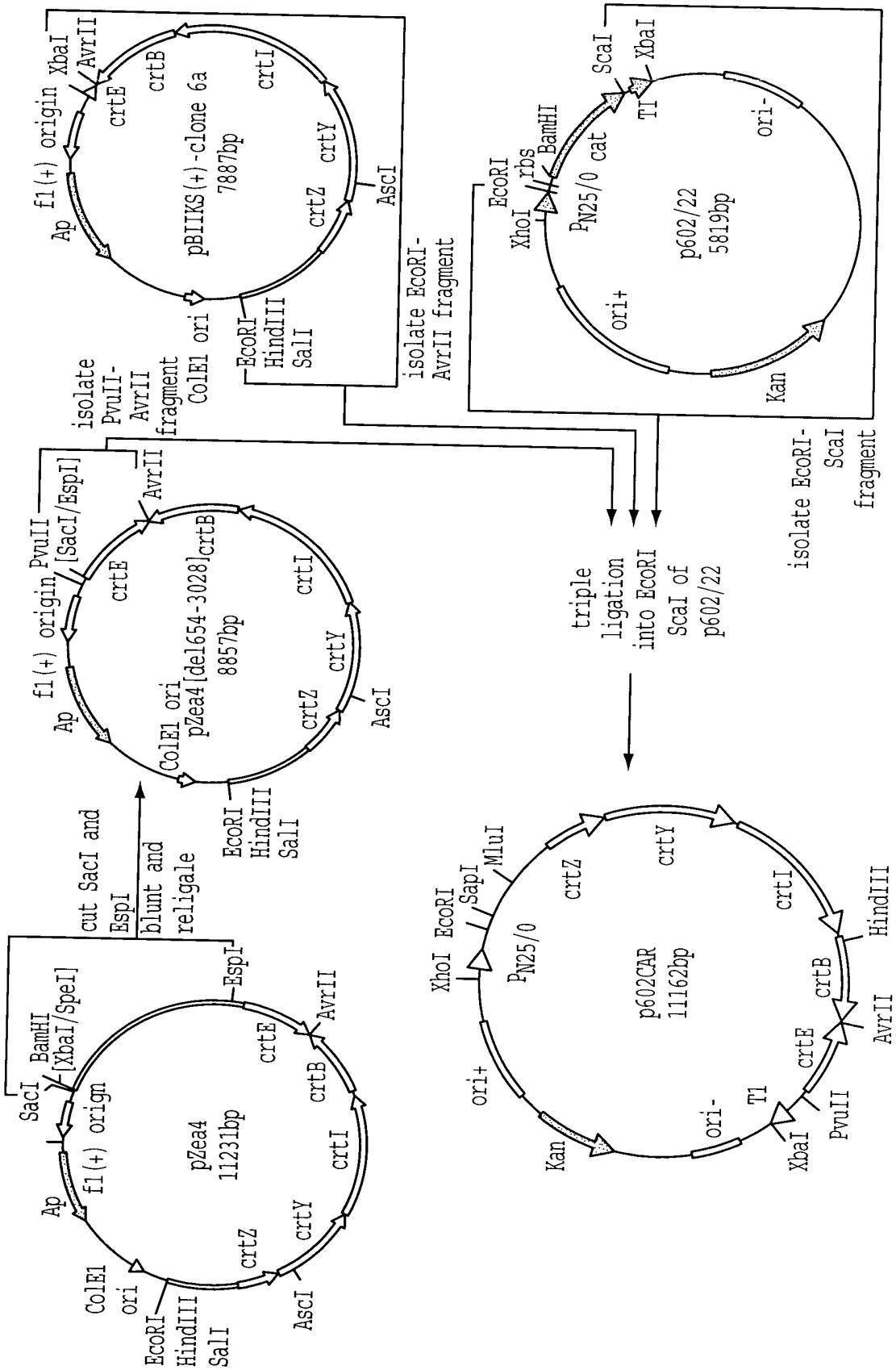

pBIIKS(+)-clone 6a
7887bp
f1(+) origin
XbaI
AvrII
crtE
crtB
crtI
crtY
crtZ
AscI
Ap
ColE1 ori
EcoRI
HindIII
SalI
isolate EcoRI-AvrII fragment
isolate PvuII-AvrII fragment
p602/22
5819bp
EcoRI
XhoI
rbs
BamHI
PN25/0 cat
ScaI
T1
ori-
ori+
Kan
isolate EcoRI-ScaI fragment
triple ligation into EcoRI ScaI of p602/22
pZea4[del654-3028]crtB
8857bp
PvuII
[SacI/EspI]
cut SacI and EspI blunt and religale
pZea4
11231bp
SacI
BamHI
[XbaI/SpeI]
EspI
p602CAR
11162bp
SapI
MluI
HindIII
PvuII

FIG. 17

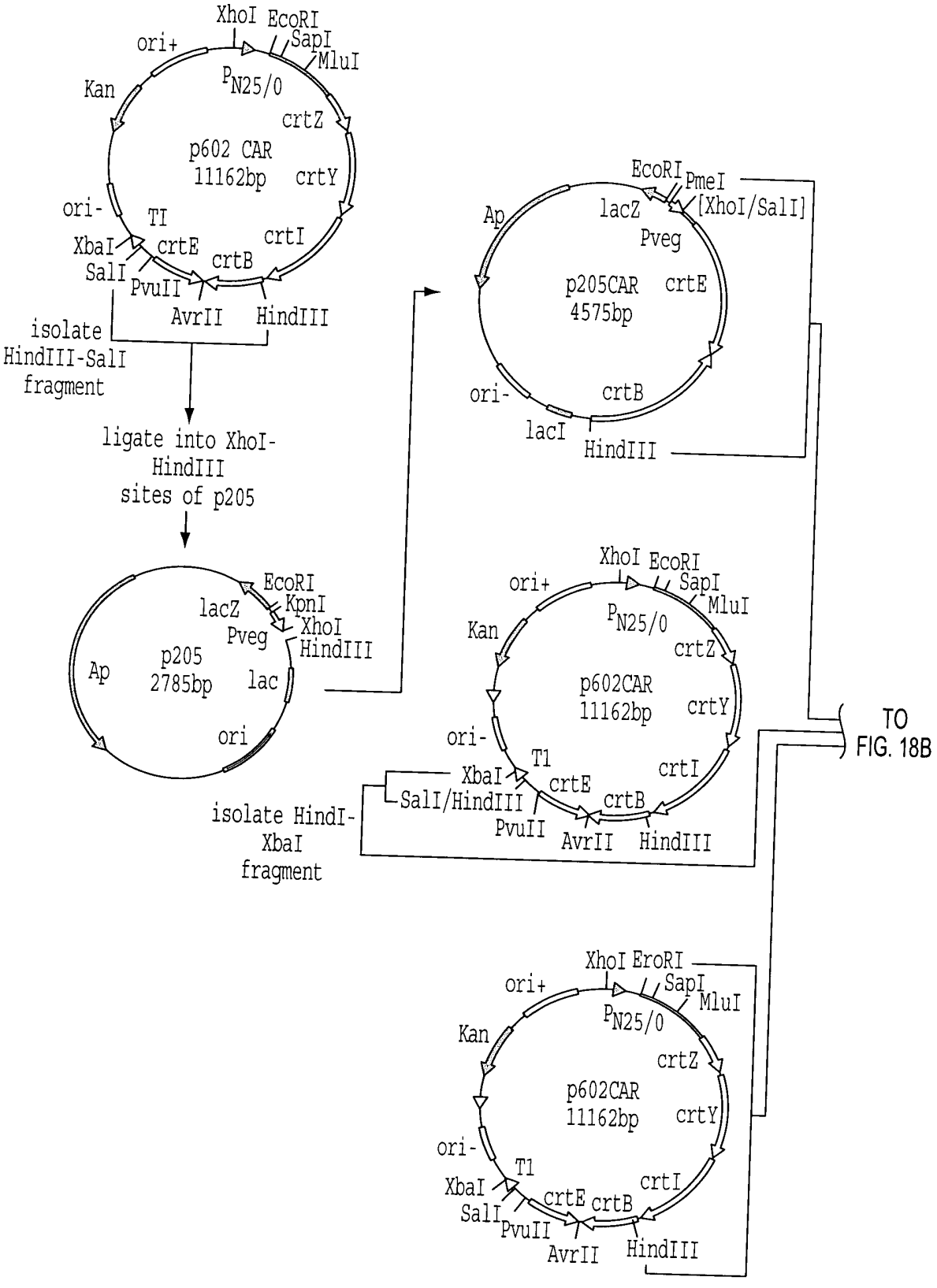

XhoI
EcoRI
SapI
MluI
ori+
Kan
$P_{N25/0}$
crtZ
p602 CAR
11162bp
crtY
ori-
TI
XbaI
crtE
crtI
SalI
crtB
PvuII
AvrII
HindIII
isolate HindIII-SalI fragment
ligate into XhoI-HindIII sites of p205
EcoRI
PmeI
[XhoI/SalI]
Ap
lacZ
Pveg
p205CAR
4575bp
crtE
ori-
crtB
lacI
HindIII
EcoRI
KpnI
XhoI
HindIII
lacZ
Pveg
p205
2785bp
Ap
lac
ori
XhoI
EcoRI
ori+
SapI
MluI
Kan
$P_{N25/0}$
crtZ
p602CAR
11162bp
crtY
ori-
T1
XbaI
crtE
crtI
SalI/HindIII
crtB
PvuII
AvrII
HindIII
isolate HindI-XbaI fragment
TO FIG. 18B
XhoI
EroRI
SapI
MluI
ori+
Kan
$P_{N25/0}$
crtZ
p602CAR
11162bp
crtY
ori-
T1
XbaI
crtI
crtE
crtB
SalI
PvuII
AvrII
HindIII

FIG. 18A

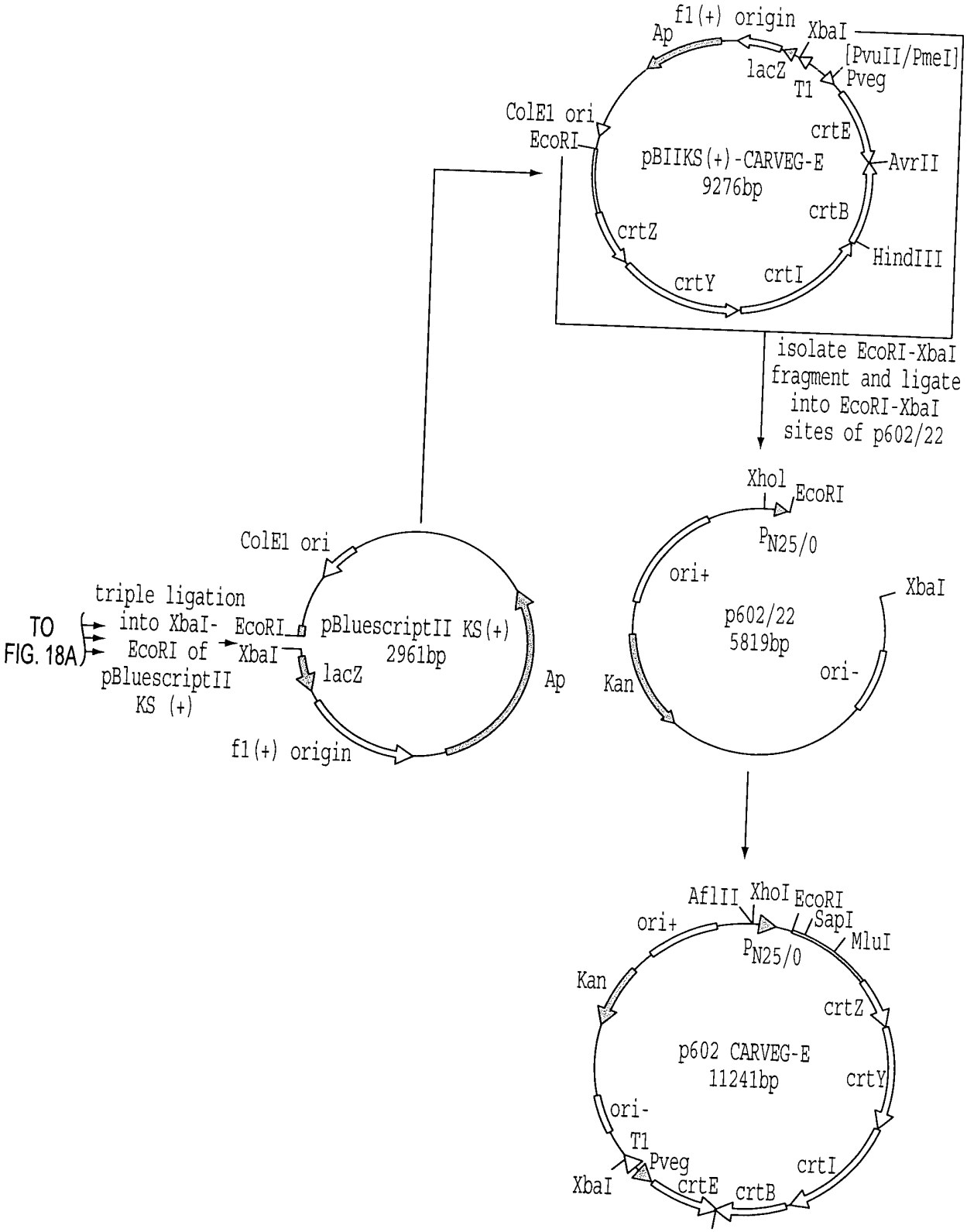

f1(+) origin
Ap
XbaI
[PvuII/PmeI]
Pveg
lacZ
T1
ColE1 ori
EcoRI
crtE
pBIIKS(+)-CARVEG-E
9276bp
AvrII
crtB
crtZ
HindIII
crtY
crtI
isolate EcoRI-XbaI fragment and ligate into EcoRI-XbaI sites of p602/22
XhoI
EcoRI
P N25/0
ori+
XbaI
p602/22
5819bp
ori-
Kan
ColE1 ori
triple ligation into XbaI-EcoRI of pBluescriptII KS (+)
TO FIG. 18A
EcoRI
XbaI
pBluescriptII KS(+)
2961bp
lacZ
Ap
f1(+) origin
AflII
XhoI
EcoRI
SapI
MluI
ori+
P N25/0
Kan
crtZ
p602 CARVEG-E
11241bp
crtY
ori-
T1
Pveg
crtI
XbaI
crtE
crtB

FIG. 18B

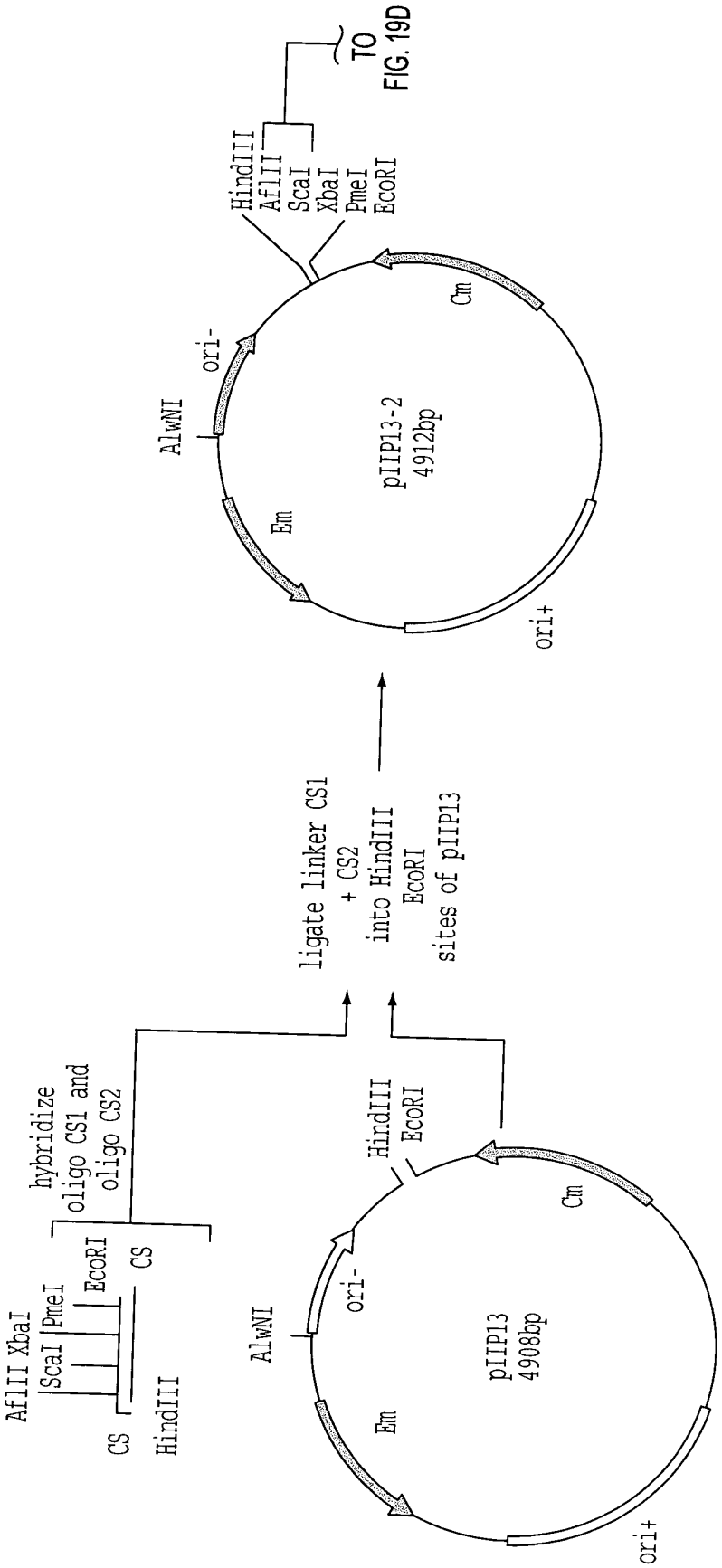

AflII XbaI
ScaI PmeI
EcoRI
CS
CS
HindIII
hybridize
oligo CS1 and
oligo CS2
ligate linker CS1
+ CS2
into HindIII
EcoRI
sites of pIIP13
AlwNI
ori-
Em
pIIP13
4908bp
ori+
Cm
HindIII
EcoRI
AlwNI
ori-
Em
pIIP13-2
4912bp
ori+
Cm
HindIII
AflII
ScaI
XbaI
PmeI
EcoRI
TO
FIG. 19D

FIG. 19A

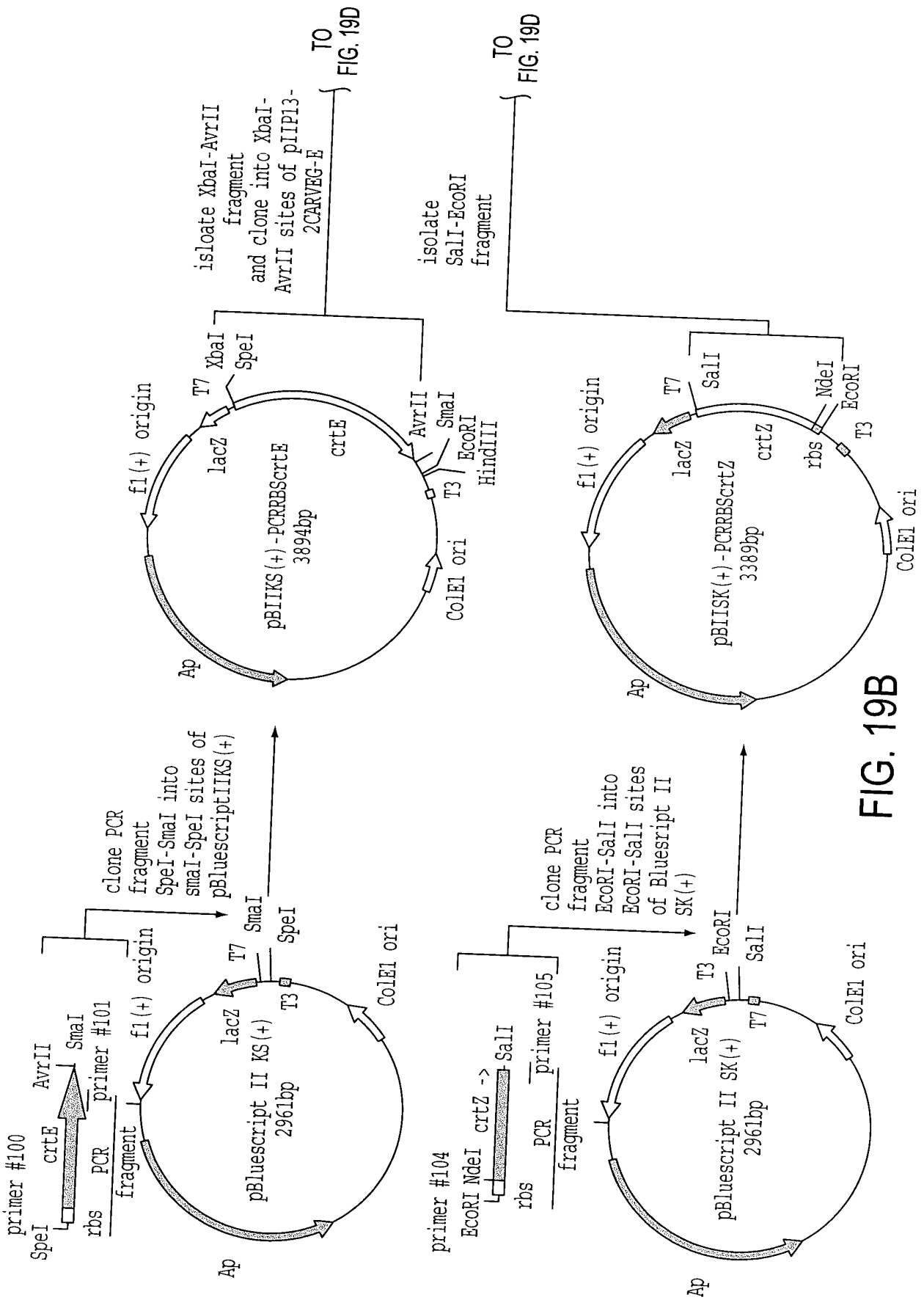
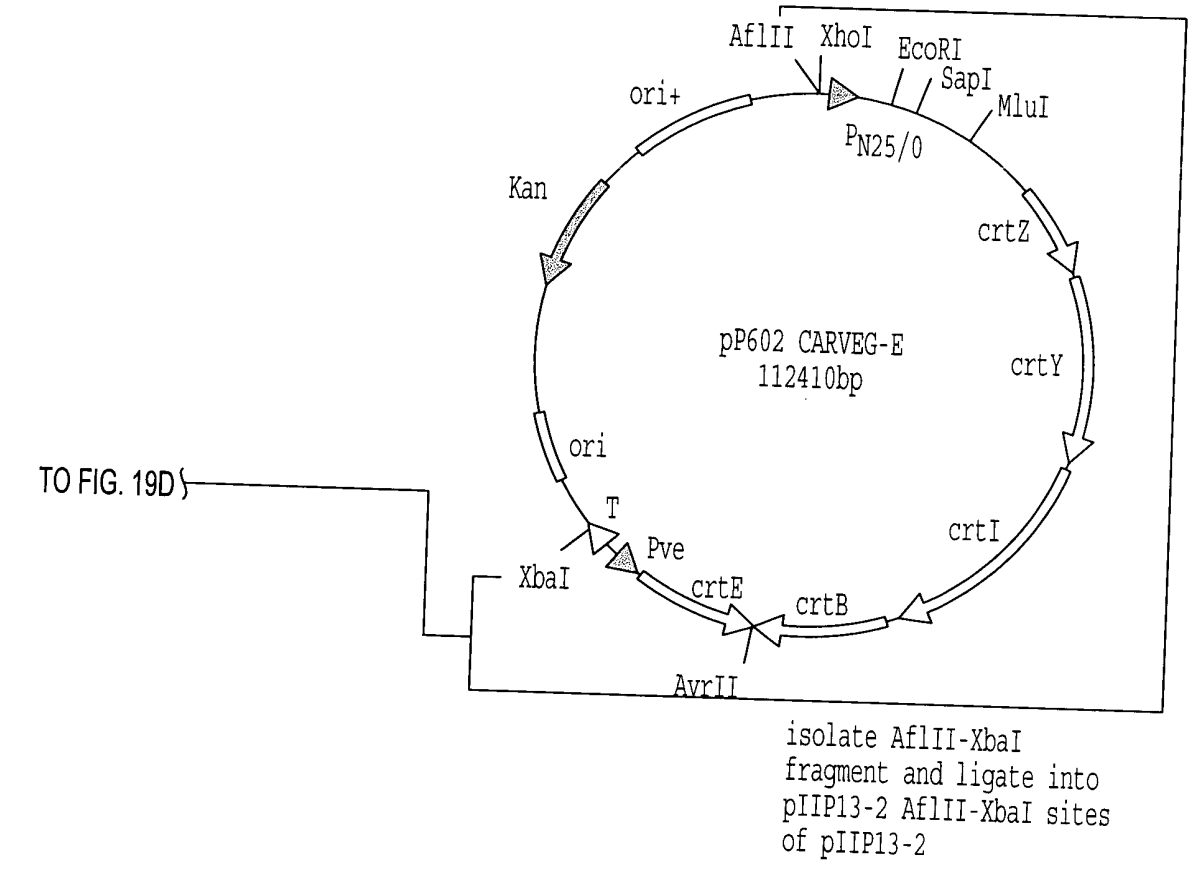

primer #100
SpeI
crtE
AvrII
SmaI
rbs
PCR fragment
primer #101
f1(+) origin
clone PCR fragment SpeI-SmaI into smaI-SpeI sites of pBluescriptIIKS(+)
pBluescript II KS(+) 2961bp
lacZ
T7
SmaI
SpeI
T3
ColE1 ori
Ap
primer #104
EcoRI NdeI crtZ ->
SalI
rbs
PCR fragment
primer #105
clone PCR fragment EcoRI-SalI into EcoRI-SalI sites of Bluesript II SK(+)
pBluescript II SK(+) 2961bp
T3
EcoRI
SalI
T7
isloate XbaI-AvrII fragment and clone into XbaI-AvrII sites of pIIP13-2CARVEG-E
TO FIG. 19D
pBIIKS(+)-PCRRBScrtE 3894bp
crtE
XbaI
SpeI
AvrII
SmaI
EcoRI
HindIII
isolate SalI-EcoRI fragment
pBIISK(+)-PCRRBScrtZ 3389bp
crtZ
SalI
NdeI
EcoRI
rbs

FIG. 19B


AflII
XhoI
EcoRI
SapI
MluI
ori+
$P_{N25/0}$
Kan
crtZ
pP602 CARVEG-E
112410bp
crtY
ori
T
crtI
Pve
XbaI
crtE
crtB
AvrII
TO FIG. 19D
isolate AflII-XbaI fragment and ligate into pIIP13-2 AflII-XbaI sites of pIIP13-2

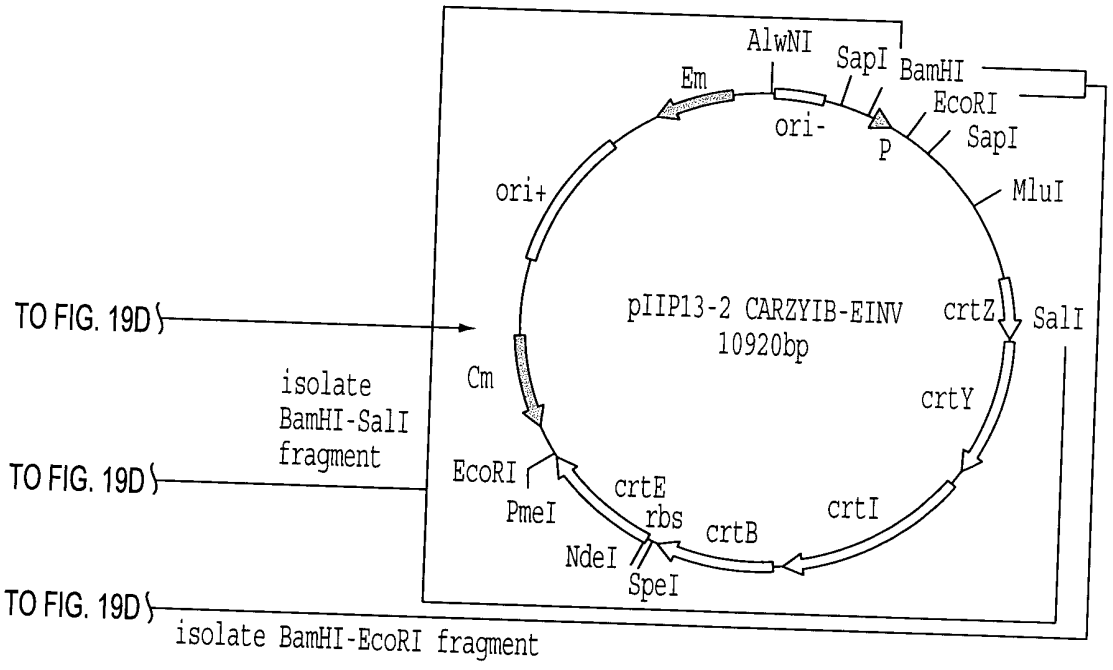

AlwNI
SapI
BamHI
Em
EcoRI
ori-
P
SapI
ori+
MluI
pIIP13-2 CARZYIB-EINV
10920bp
crtZ
SalI
TO FIG. 19D
Cm
isolate BamHI-SalI fragment
crtY
TO FIG. 19D
EcoRI
crtE
PmeI
rbs
crtI
crtB
NdeI
SpeI
TO FIG. 19D
isolate BamHI-EcoRI fragment

FIG. 19C

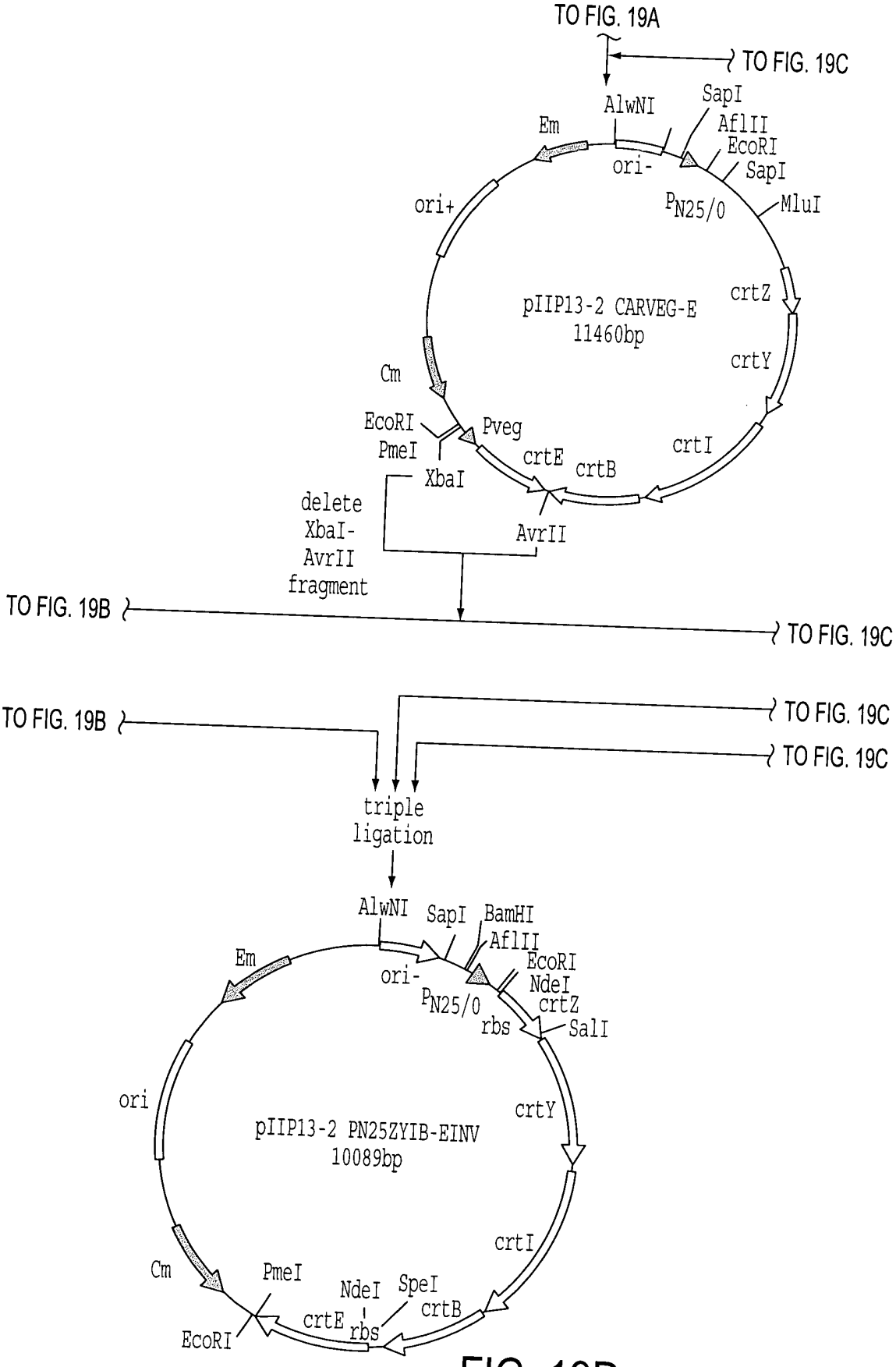

TO FIG. 19A
TO FIG. 19C
AlwNI
SapI
AflII
EcoRI
SapI
Em
ori-
ori+
$P_{N25/0}$
MluI
crtZ
pIIP13-2 CARVEG-E
11460bp
crtY
Cm
EcoRI
PmeI
Pveg
crtE
crtB
crtI
XbaI
delete
XbaI-
AvrII
fragment
AvrII
TO FIG. 19B
TO FIG. 19C
TO FIG. 19B
TO FIG. 19C
TO FIG. 19C
triple
ligation
AlwNI
SapI
BamHI
AflII
EcoRI
NdeI
Em
ori-
$P_{N25/0}$
crtZ
rbs
SalI
ori
crtY
pIIP13-2 PN25ZYIB-EINV
10089bp
crtI
Cm
PmeI
NdeI
SpeI
crtB
crtE
rbs
EcoRI

FIG. 19D

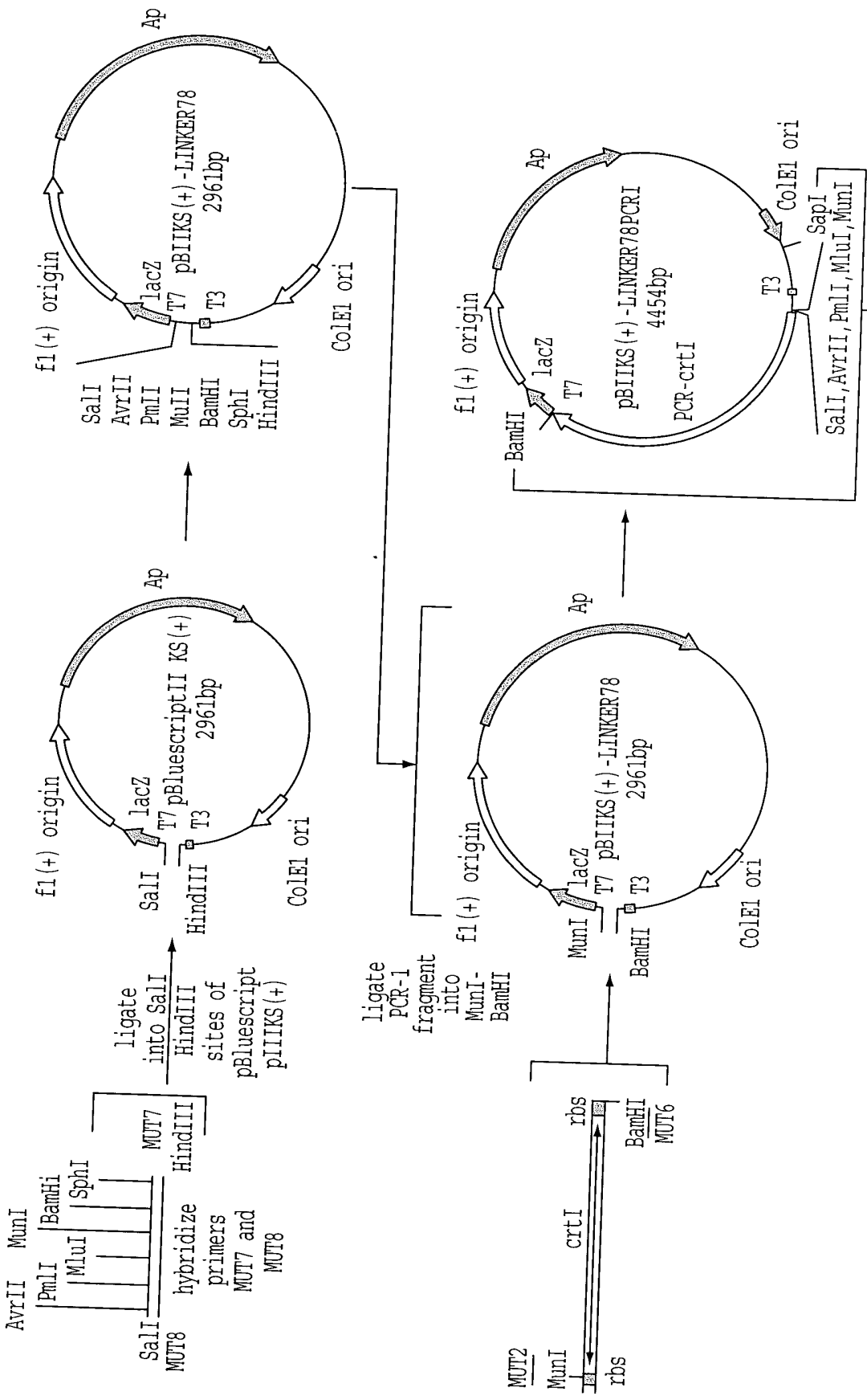

AvrII
MunI
PmlI
BamHi
MluI
SphI
SalI
MUT8
hybridize primers MUT7 and MUT8
MUT7
HindIII
ligate into SalI HindIII sites of pBluescript pIIKS(+)
f1(+) origin
lacZ
T7 pBluescriptII KS(+)
2961bp
T3
Ap
SalI
HindIII
ColE1 ori
MUT2
MunI
crtI
rbs
BamHI
MUT6
ligate PCR-1 fragment into MunI-BamHI
f1(+) origin
lacZ
T7 pBIIKS(+)-LINKER78
2961bp
T3
Ap
MunI
BamHI
ColE1 ori
SalI
AvrII
PmlI
MluI
BamHI
SphI
HindIII
pBIIKS(+)-LINKER78PCRI
4454bp
PCR-crtI
SapI
SalI,AvrII,PmlI,MluI,MunI
TO FIG. 20A-2

FIG. 20A-1

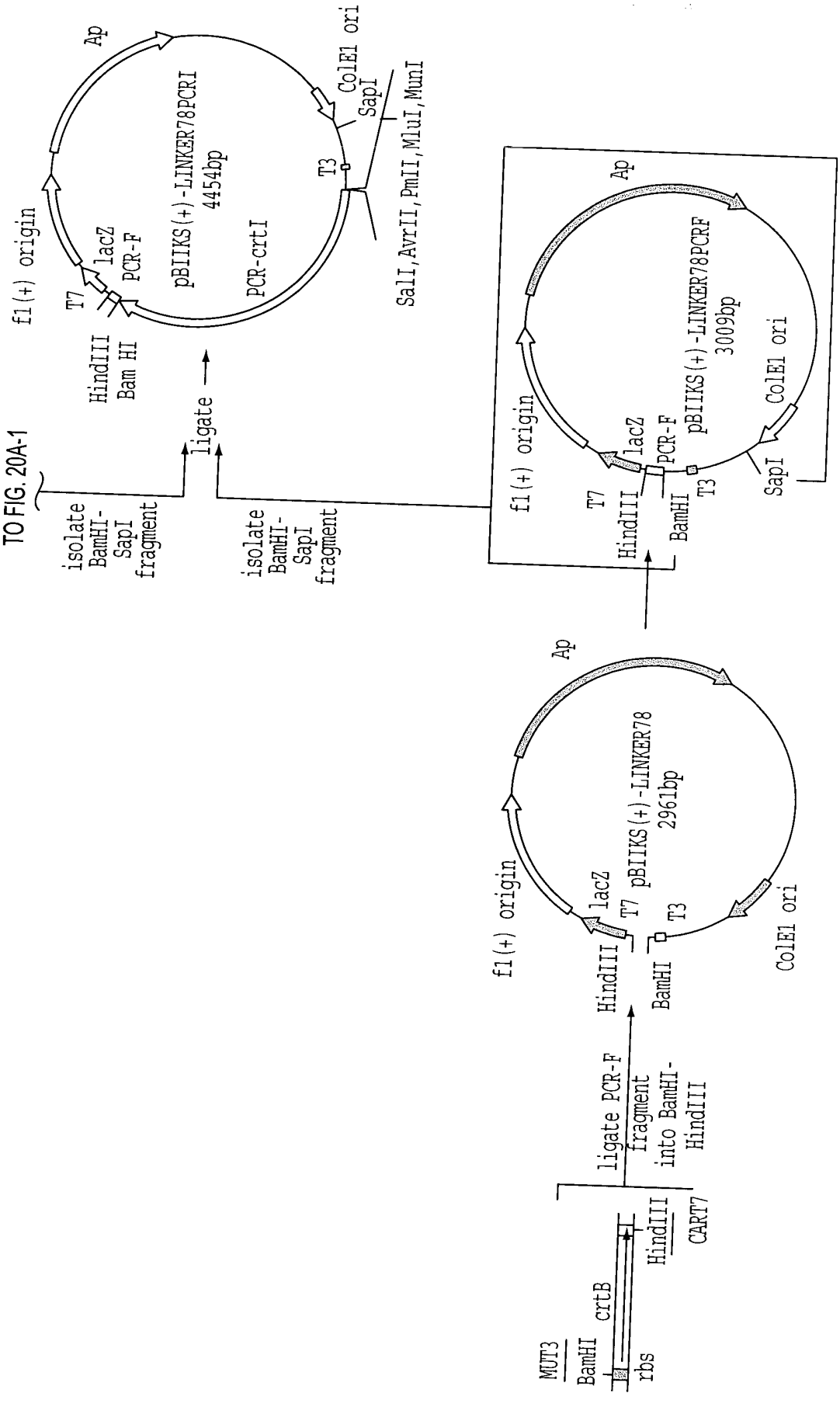

Ap
ColE1 ori
SapI
pBIIKS(+)-LINKER78PCRI
4454bp
T3
f1(+) origin
lacZ
PCR-F
T7
PCR-crtI
SalI, AvrII, PmlI, MluI, MunI
HindIII
Bam HI
TO FIG. 20A-1
isolate BamHI-SapI fragment
ligate
isolate BamHI-SapI fragment
Ap
pBIIKS(+)-LINKER78PCRF
3009bp
ColE1 ori
f1(+) origin
lacZ
PCR-F
T7
T3
HindIII
BamHI
SapI
Ap
pBIIKS(+)-LINKER78
2961bp
f1(+) origin
lacZ
T7
T3
HindIII
BamHI
ColE1 ori
ligate PCR-F fragment into BamHI-HindIII
CART7
MUT3
BamHI
crtB
rbs
HindIII

FIG. 20A-2

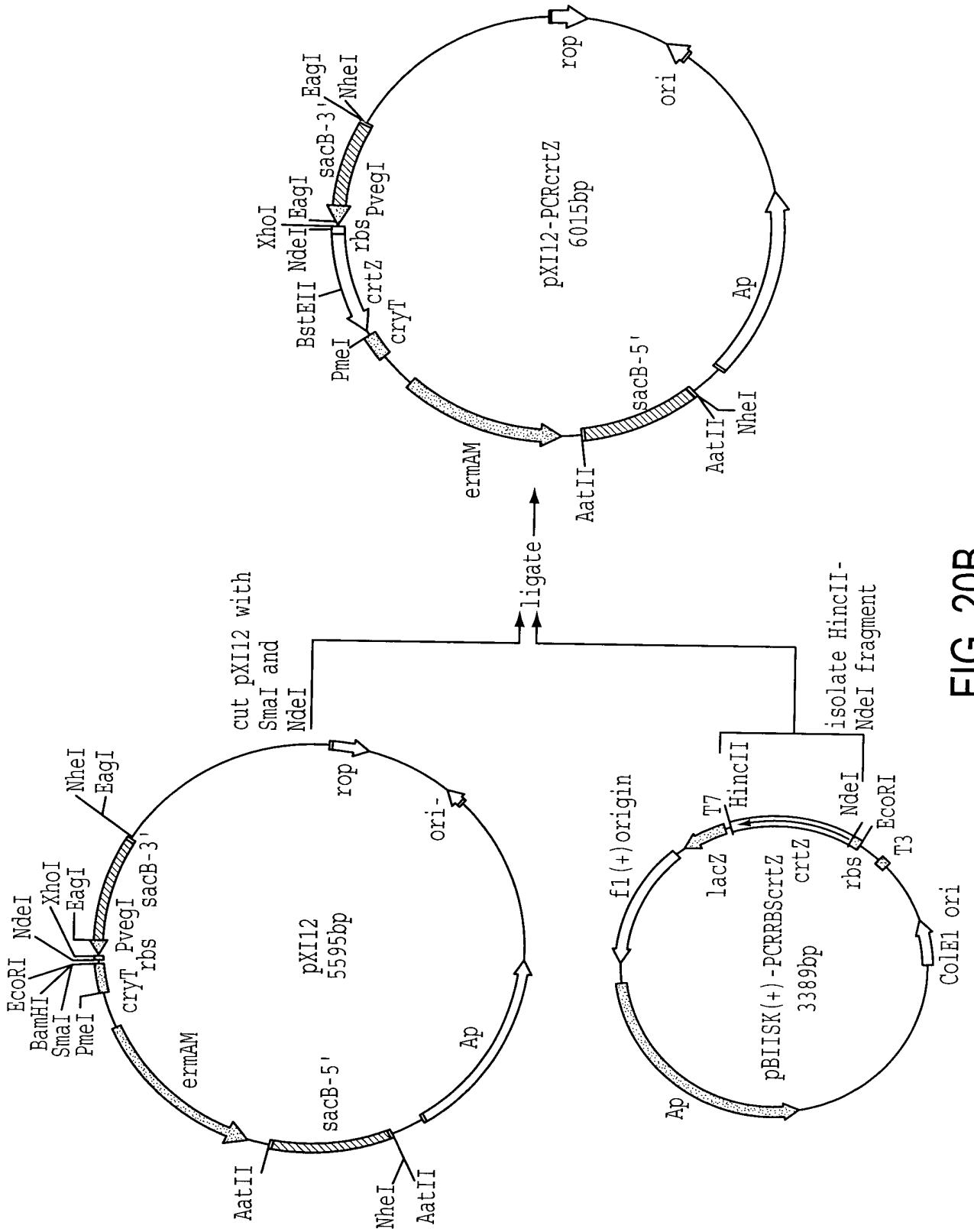

pXI12-PCRcrtZ
6015bp
rop
ori
Ap
AatII
NheI
ermAM
cryT
crtZ
rbs
PvegI
sacB-3'
sacB-5'
EagI
NheI
XhoI
NdeI
BstEII
PmeI
ligate
cut pXI12 with
SmaI and
NdeI
pXI12
5595bp
ori-
EcoRI
BamHI
SmaI
PmeI
XhoI
NdeI
isolate HincII-
NdeI fragment
f1(+)origin
T7
HincII
lacZ
pBIISK(+)-PCRRBScrtZ
3389bp
T3
EcoRI
ColE1 ori

FIG. 20B

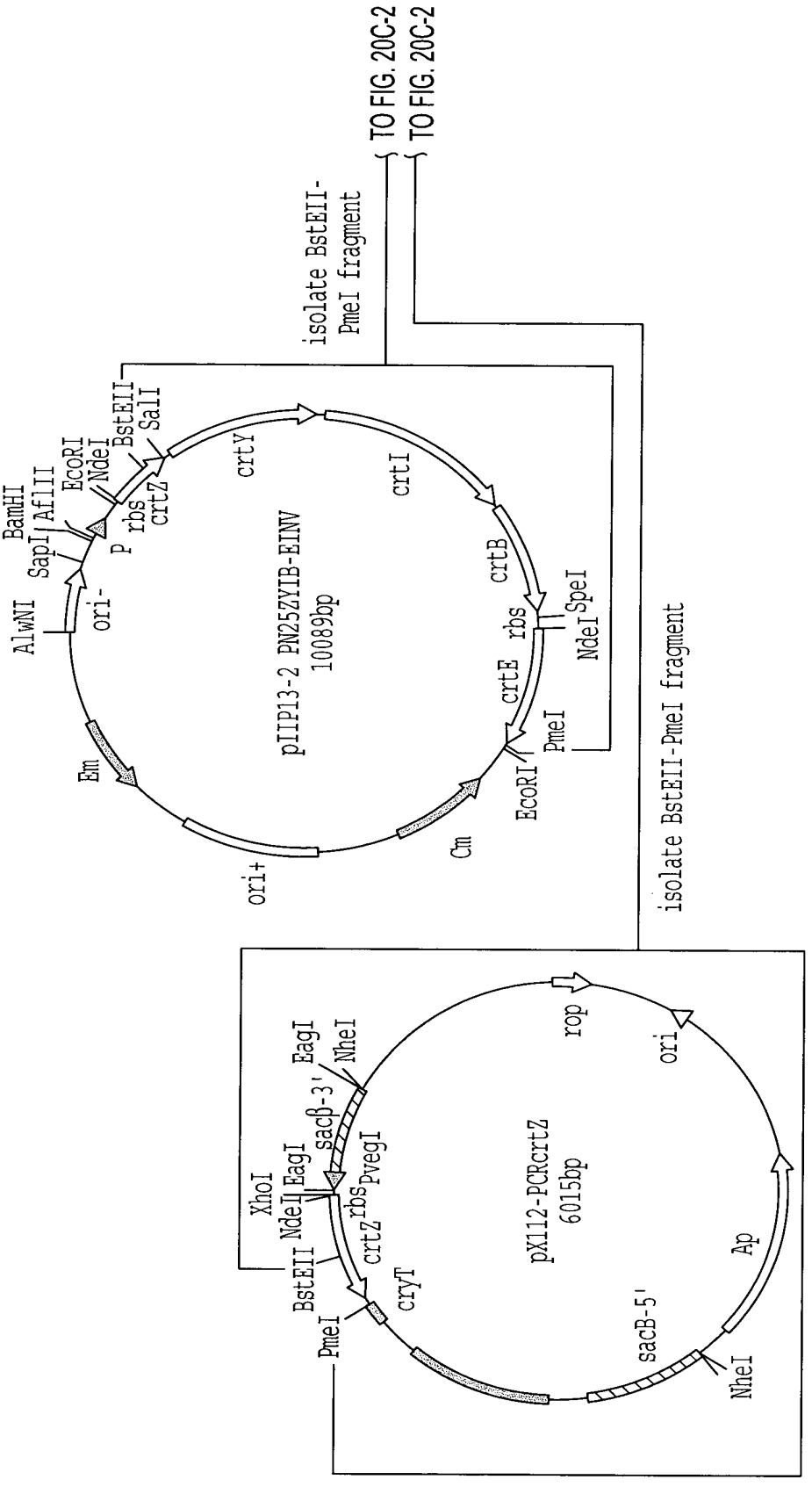

TO FIG. 20C-2
TO FIG. 20C-2
isolate BstEII-PmeI fragment
pIIP13-2 PN25YIB-EINV
10089bp
crtY
crtI
crtB
crtE
rbs
NdeI
SpeI
PmeI
EcoRI
Cm
ori+
Em
ori-
AlwNI
SapI
BamHI
AflII
EcoRI
NdeI
BstEII
SalI
rbs
crtZ
P
isolate BstEII-PmeI fragment
pX112-PCRcrtZ
6015bp
rop
ori
Ap
NheI
sacB-5'
cryT
crtZ
rbs
PvegI
sacβ-3'
EagI
NheI
XhoI
EagI
NdeI
BstEII
PmeI

FIG. 20C-1

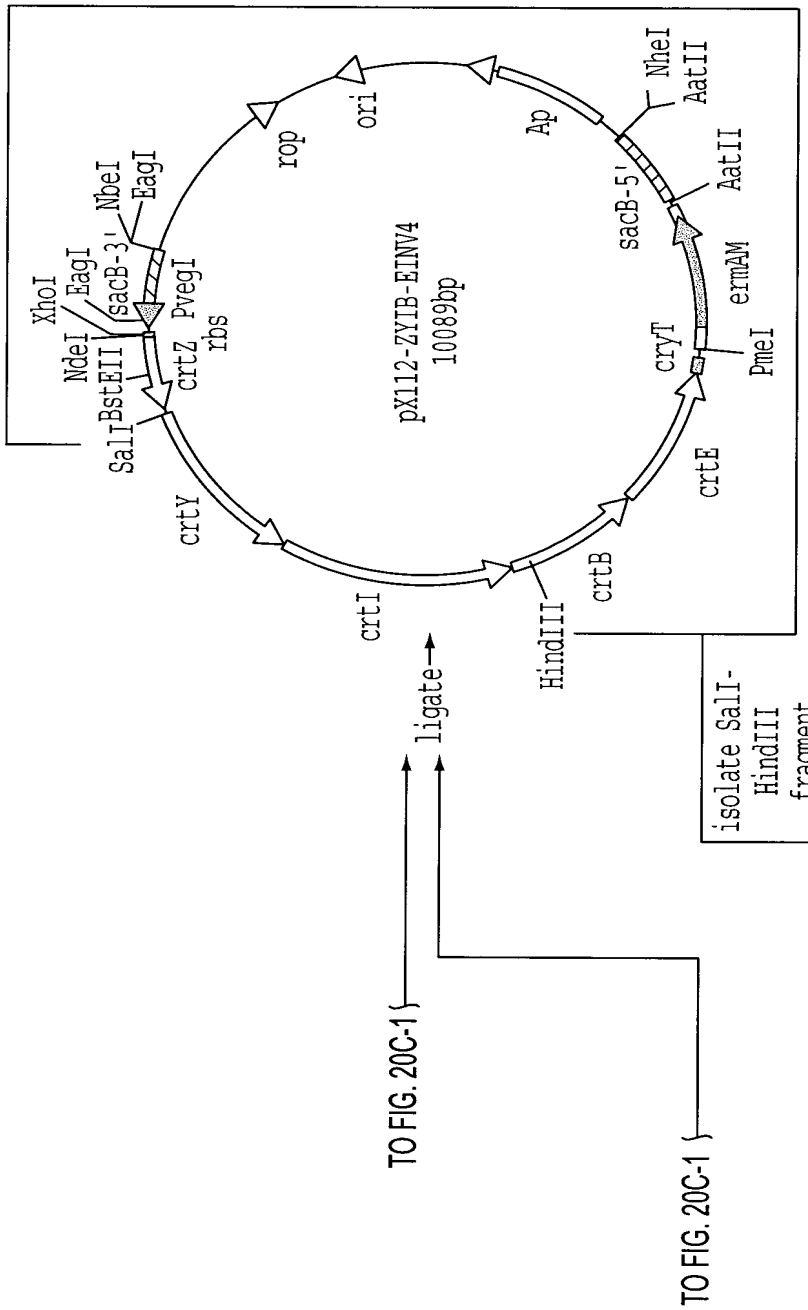

pX112-ZYIB-EINV4
10089bp
rop
ori
Ap
NheI
AatII
AatII
sacB-5'
ermAM
PmeI
cryT
crtE
crtB
HindIII
crtI
crtY
SalI
BstEII
NdeI
XhoI
EagI
sacB-3'
NheI
EagI
PvegI
rbs
crtZ
ligate
isolate SalI-HindIII fragment
TO FIG. 20C-1
TO FIG. 20C-1
TO FIG. 20C-4

FIG. 20C-2

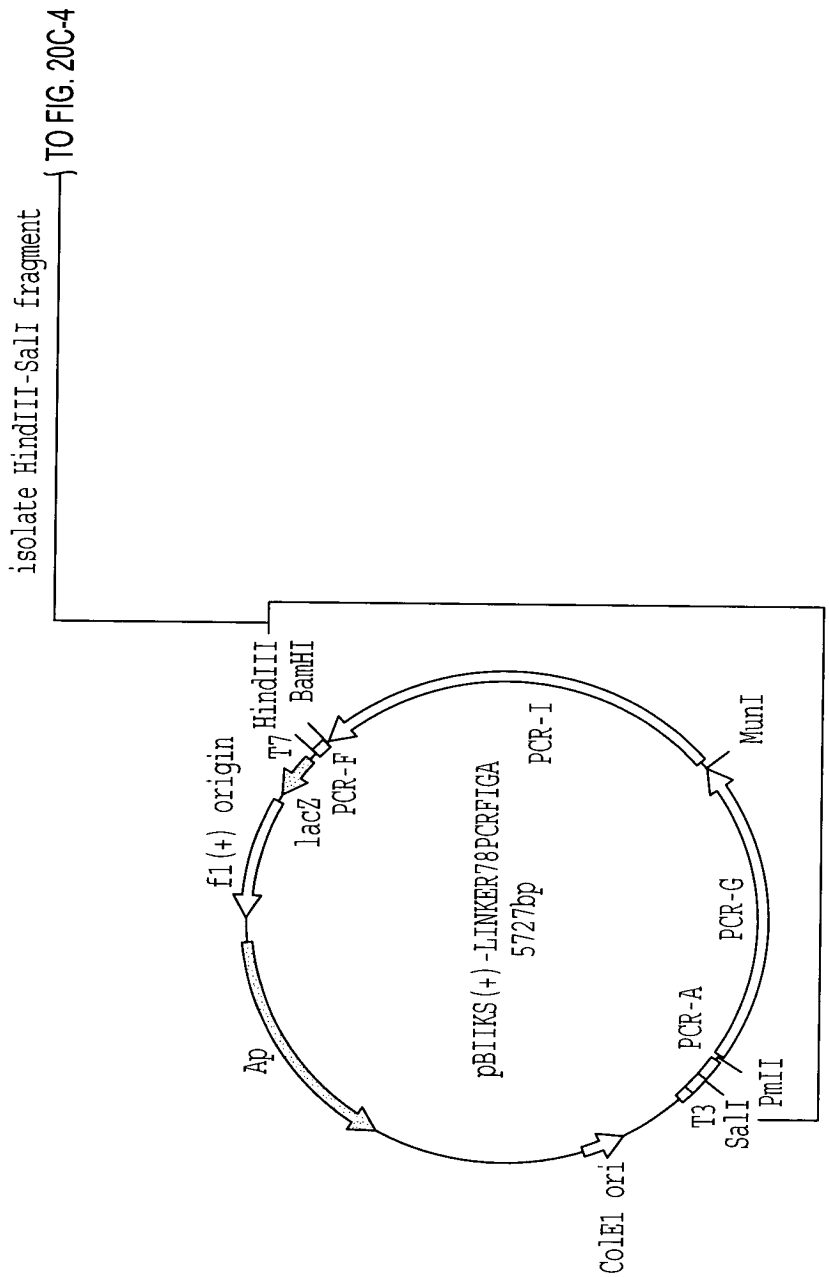

isolate HindIII-SalI fragment
TO FIG. 20C-4
HindIII
BamHI
T7
PCR-F
lacZ
f1(+) origin
Ap
ColE1 ori
T3
SalI
PmlI
PCR-A
PCR-G
MunI
PCR-I
pBIIKS(+)-LINKER78PCRFTGA
5727bp

FIG. 20C-3

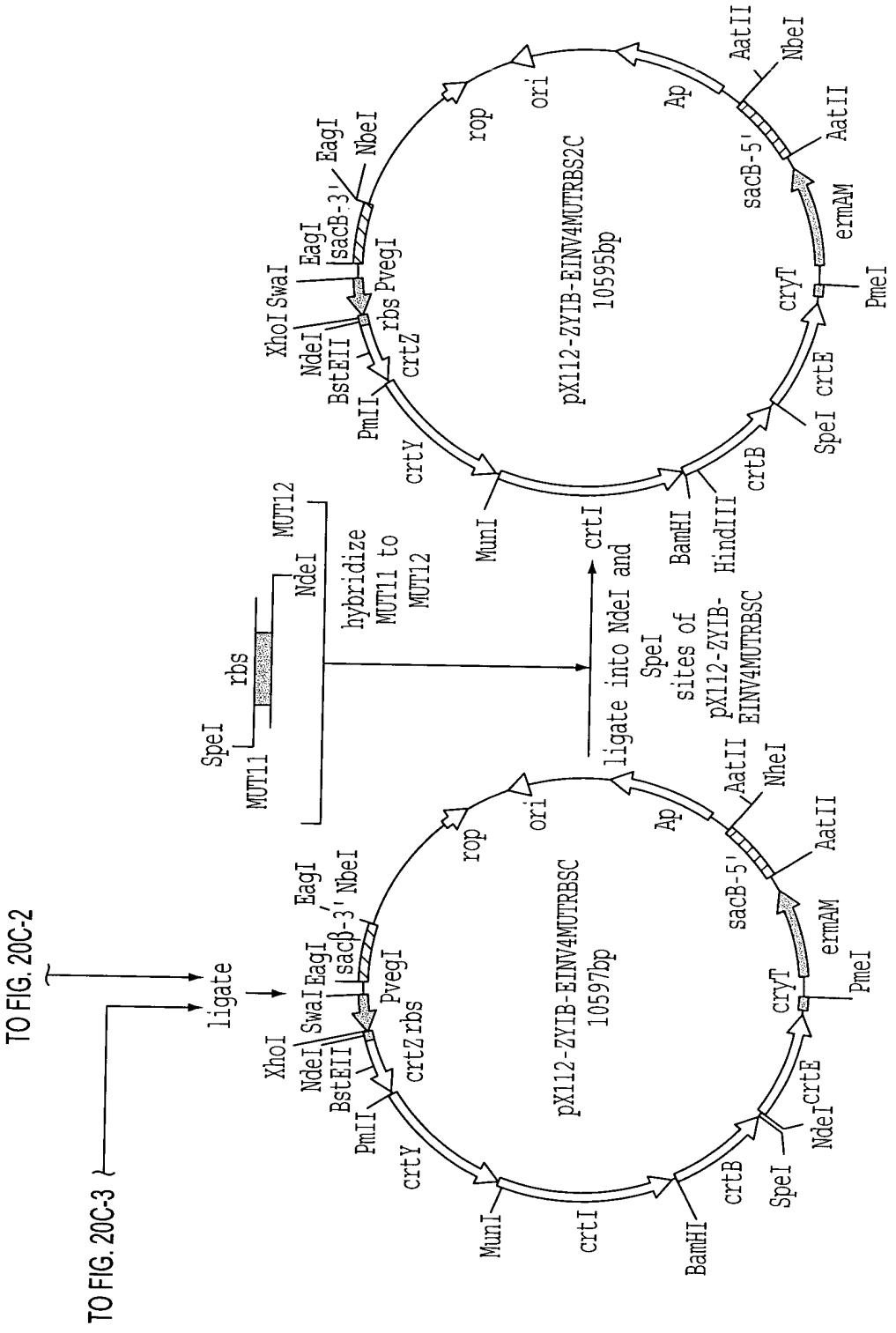

TO FIG. 20C-2
TO FIG. 20C-3
ligate
XhoI
SwaI
NdeI
BstEII
PmlI
crtZ
rbs
PvegI
EagI
sacB-3'
NbeI
crtY
MunI
crtI
BamHI
crtB
SpeI
NdeI
crtE
cryT
PmeI
ermAM
AatII
NheI
Ap
ori
rop
pX112-ZYIB-EINV4MUTRBSC
10597bp
MUT11
SpeI
rbs
NdeI
MUT12
hybridize MUT11 to MUT12
ligate into NdeI and SpeI sites of pX112-ZYIB-EINV4MUTRBSC
HindIII
pX112-ZYIB-EINV4MUTRBS2C
10595bp

FIG. 20C-4

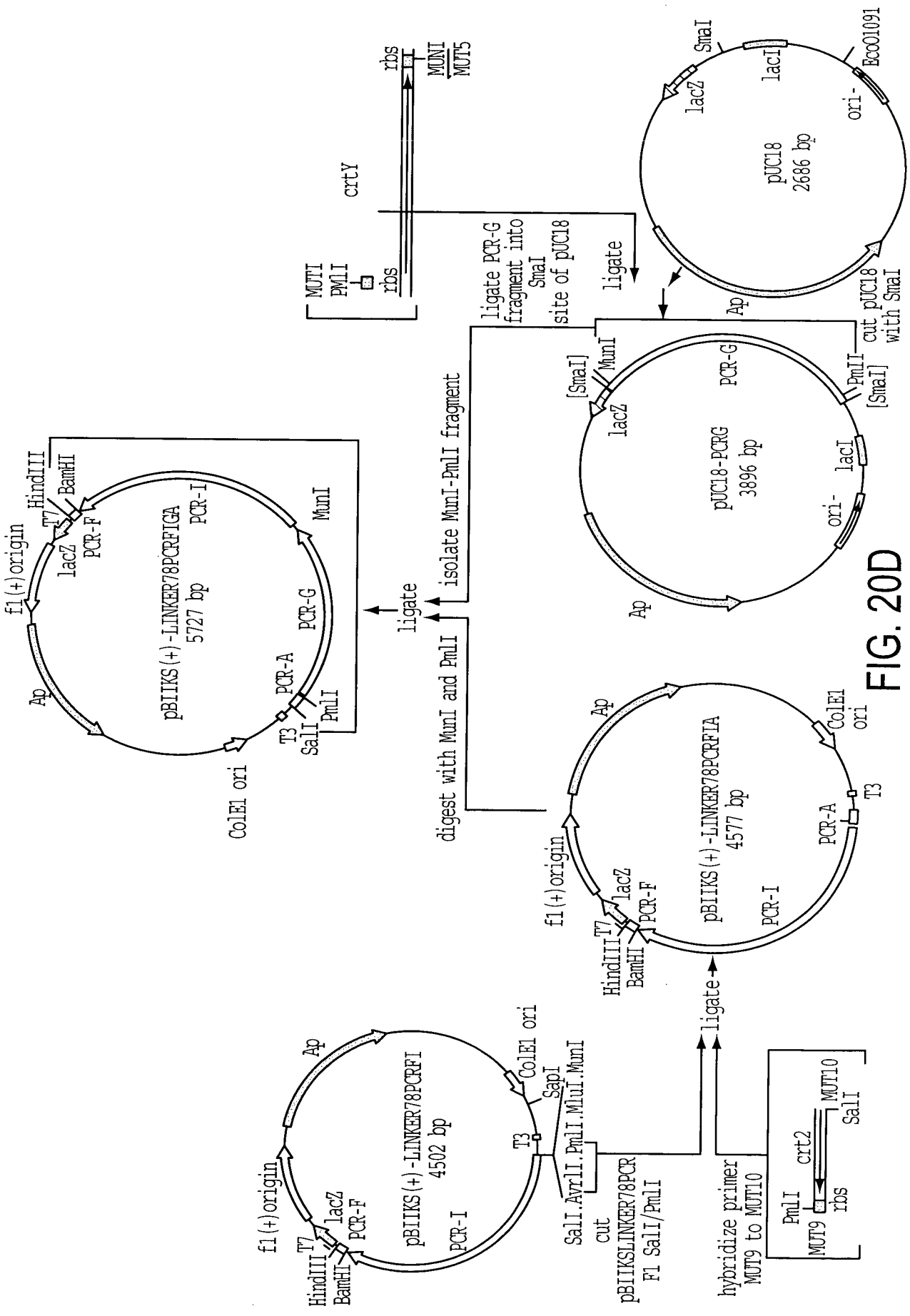

MUT1
PMLI
rbs
crtY
rbs
MUNI
MUT5
ligate PCR-G fragment into SmaI site of pUC18
ligate
pUC18
2686 bp
lacZ
SmaI
lacI
ori-
Eco01091
Ap
cut pUC18 with SmaI
[SmaI]
MunI
lacZ
pUC18-PCRG
3896 bp
PCR-G
PmlI
[SmaI]
ori-
lacI
Ap
ligate
isolate MunI-PmlI fragment
digest with MunI and PmlI
f1(+)origin
HindIII
T7
BamHI
lacZ
PCR-F
pBIIKS(+)-LINKER78PCRFIGA
5727 bp
PCR-I
MunI
PCR-G
PCR-A
PmlI
T3
SalI
Ap
ColE1 ori
f1(+)origin
HindIII
T7
BamHI
lacZ
PCR-F
pBIIKS(+)-LINKER78PCRFI
4502 bp
PCR-I
T3
ColE1 ori
SapI
SalI.AvrII.PmlI.MluI.MunI
cut
pBIIKSLINKER78PCR
F1 SalI/PmlI
ligate
f1(+)origin
HindIII
T7
BamHI
lacZ
PCR-F
pBIIKS(+)-LINKER78PCRFIA
4577 bp
PCR-I
Ap
ColE1
ori
T3
PCR-A
hybridize primer
MUT9 to MUT10
PmlI
crt2
MUT9
rbs
MUT10
SalI

FIG. 20D

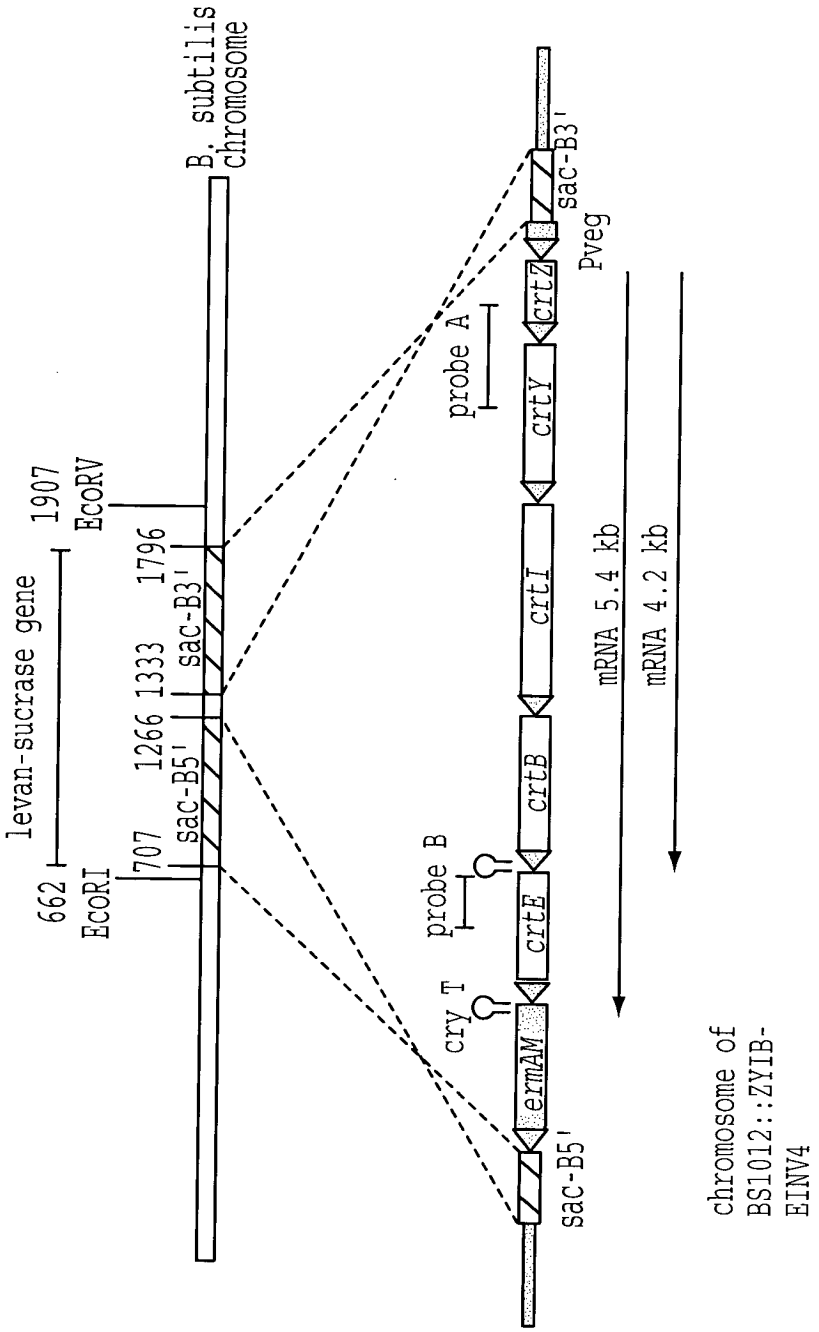
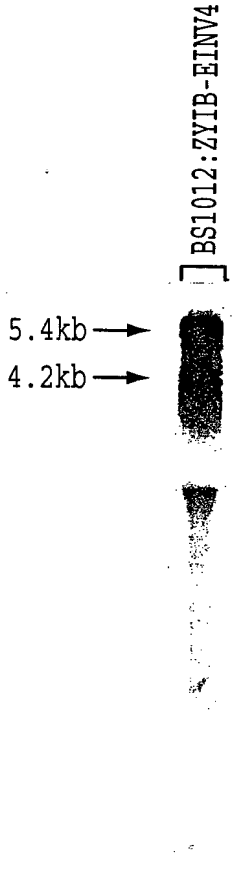

B. subtilis chromosome
levan-sucrase gene
662
EcoRI
707
sac-B5'
1266
1333
sac-B3'
1796
1907
EcoRV
probe B
cry T
ermAM
crtE
crtB
crtI
crtY
crtZ
Pveg
sac-B3'
probe A
mRNA 5.4 kb
mRNA 4.2 kb
sac-B5'
chromosome of BS1012::ZYIB-EINV4

FIG. 21A



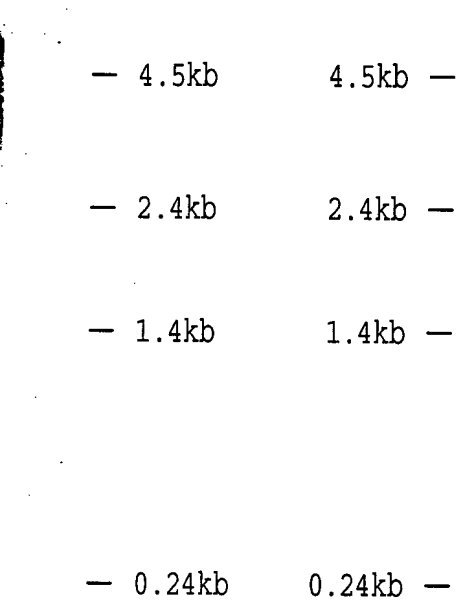


FIG. 21B

FIG. 21C

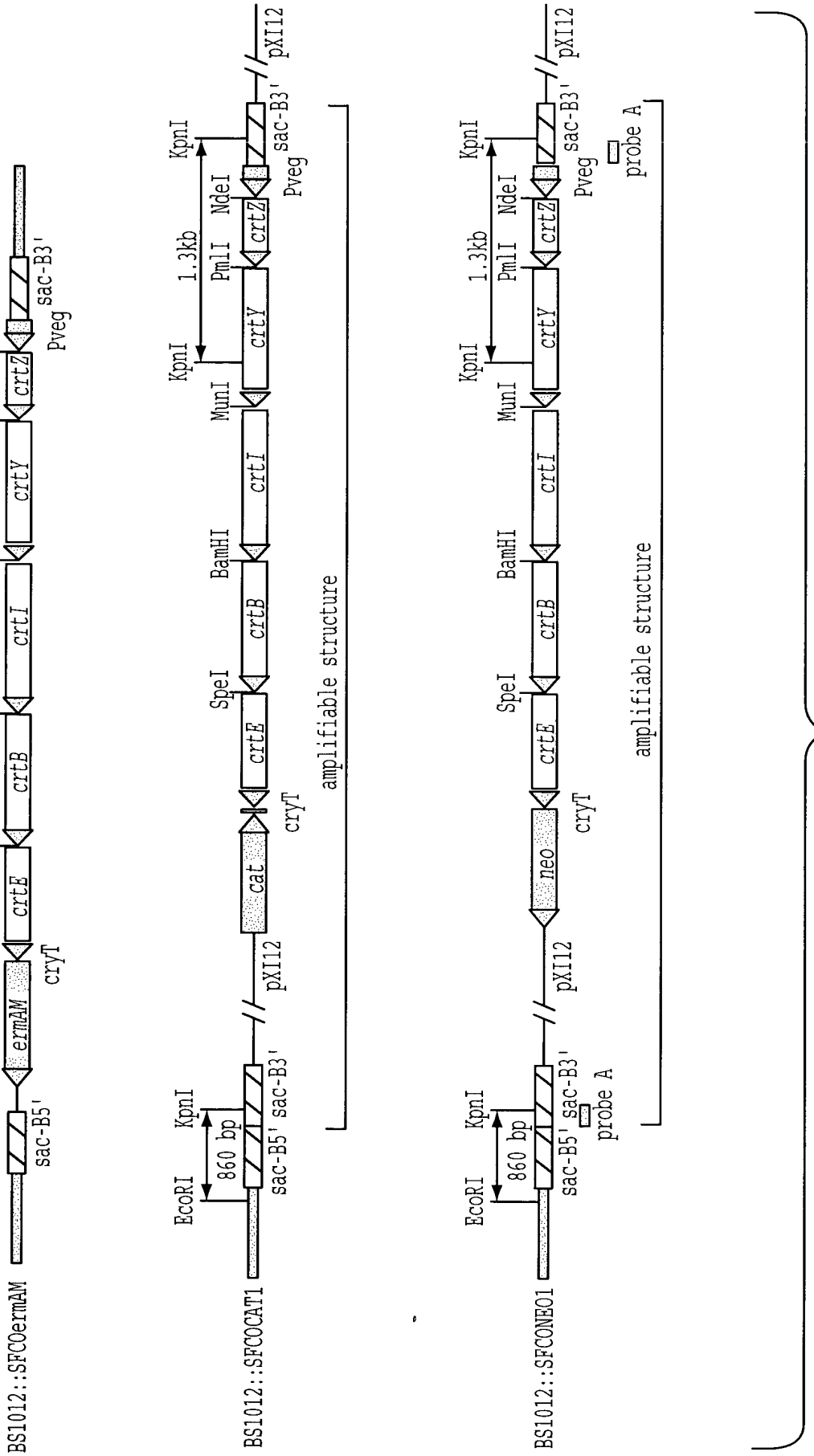

BS1012::SFCOermAM
sac-B5'
ermAM
cryT
SpeI
crtE
BamHI
crtB
MunI
crtI
PmlI
crtY
NdeI
crtZ
Pveg
sac-B3'
BS1012::SFCOCAT1
EcoRI
KpnI
860 bp
cat
pXI12
amplifiable structure
1.3kb
BS1012::SFCONEO1
neo
probe A

FIG. 22

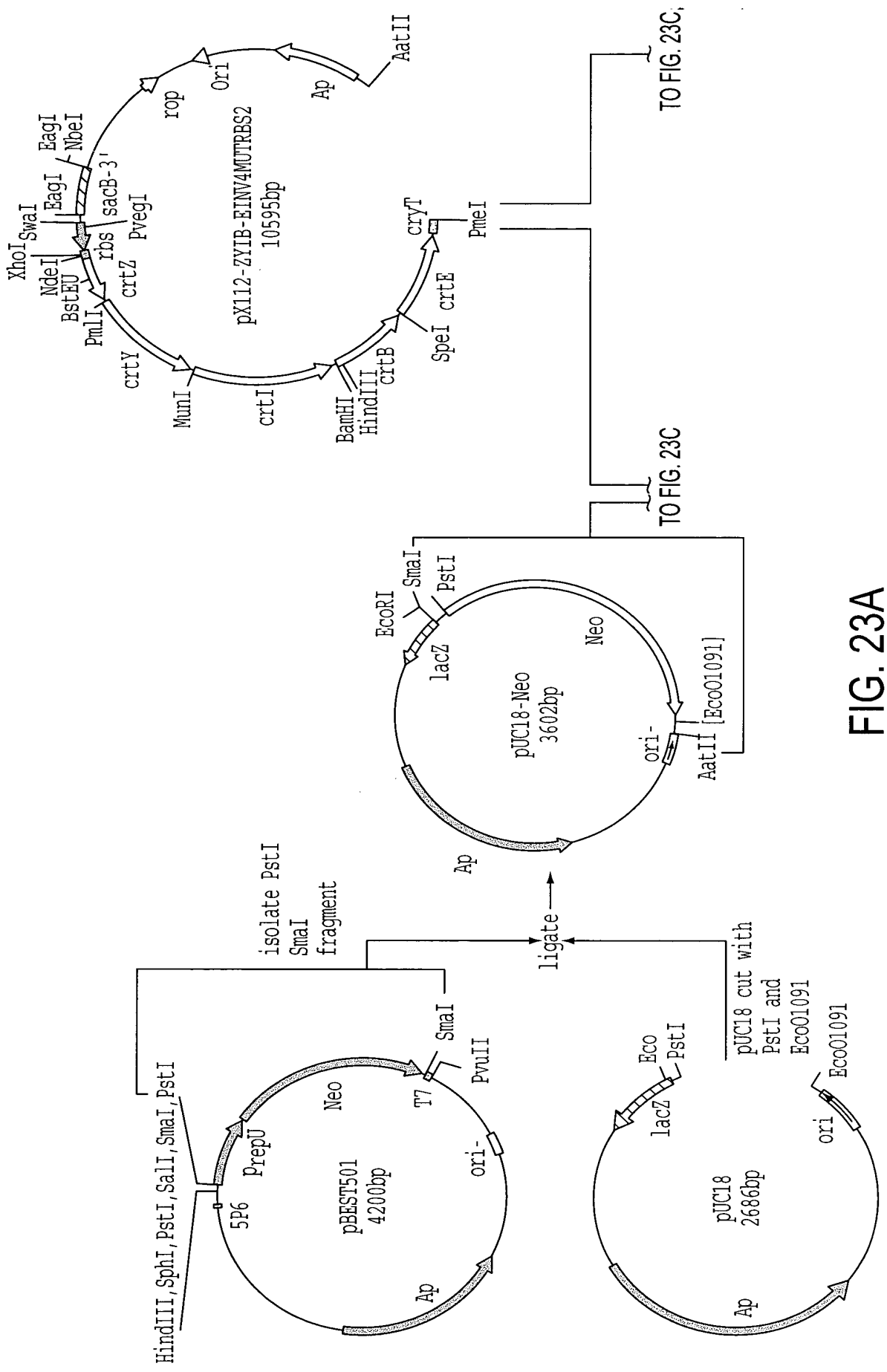

pX112-ZYIB-EINV4MUTRBS2
10595bp
AatII
Ap
Ori
rop
EagI
NbeI
sacB-3'
PvegI
rbs
crtZ
crtY
MunI
crtI
BamHI
HindIII
crtB
SpeI
crtE
crtT
PmeI
XhoI
SwaI
EagI
NdeI
BstEU
PmlI
TO FIG. 23C
TO FIG. 23C
pUC18-Neo
3602bp
EcoRI
SmaI
PstI
lacZ
Neo
ori-
AatII [Eco01091]
Ap
isolate PstI
SmaI
fragment
ligate
pBEST501
4200bp
HindIII,SphI,PstI,SalI,SmaI,PstI
5P6
PrepU
Neo
T7
SmaI
PvuII
ori-
Ap
pUC18
2686bp
pUC18 cut with
PstI and
Eco01091
Eco
PstI
lacZ
ori
Eco01091
Ap

FIG. 23A

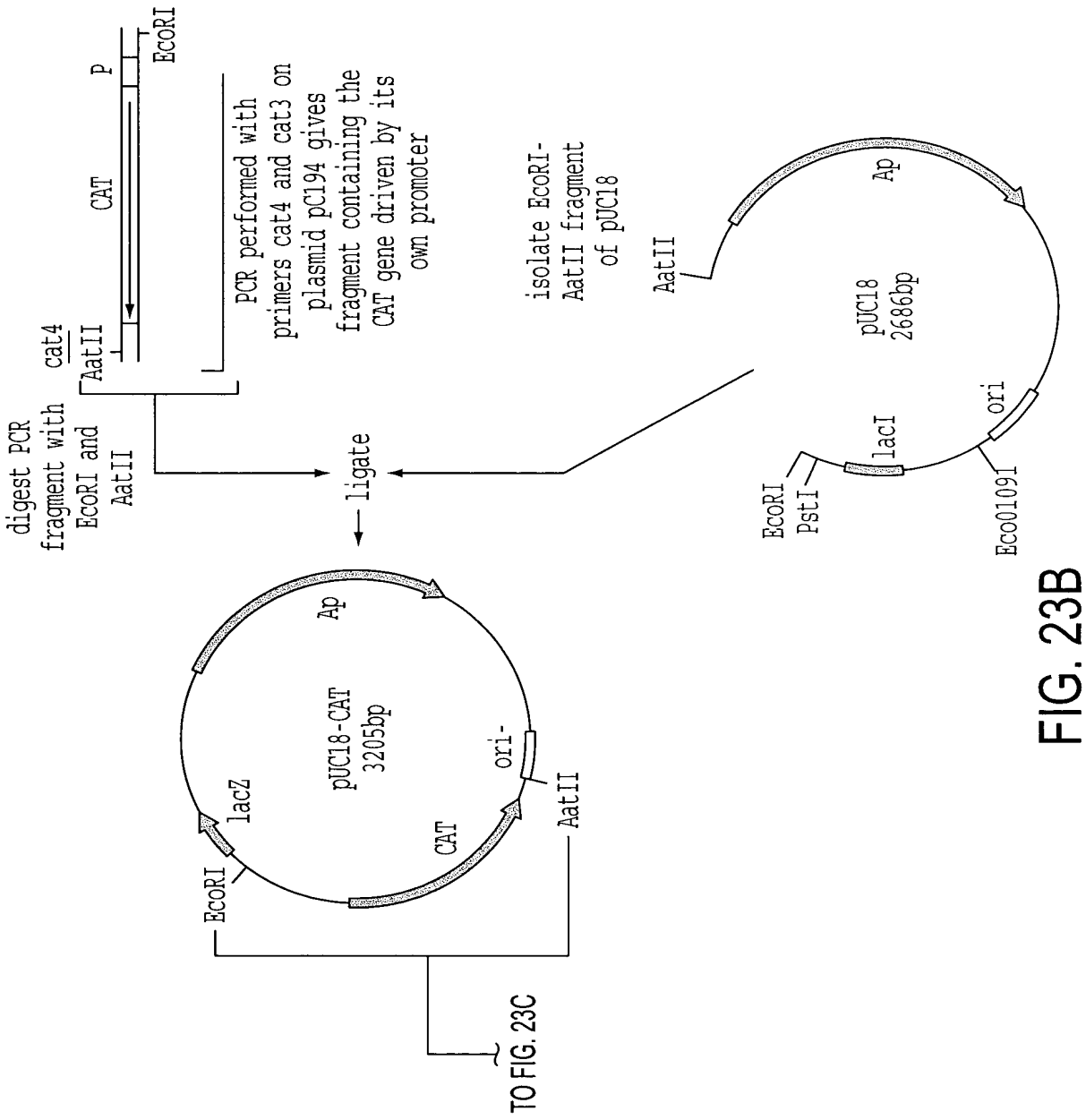

digest PCR fragment with EcoRI and AatII
cat4
AatII
CAT
P
EcoRI
PCR performed with primers cat4 and cat3 on plasmid pC194 gives fragment containing the CAT gene driven by its own promoter
isolate EcoRI-AatII fragment of pUC18
AatII
Ap
pUC18 2686bp
lacI
ori
EcoRI
PstI
Eco0109I
ligate
Ap
pUC18-CAT 3205bp
lacZ
ori-
CAT
EcoRI
AatII
TO FIG. 23C

FIG. 23B

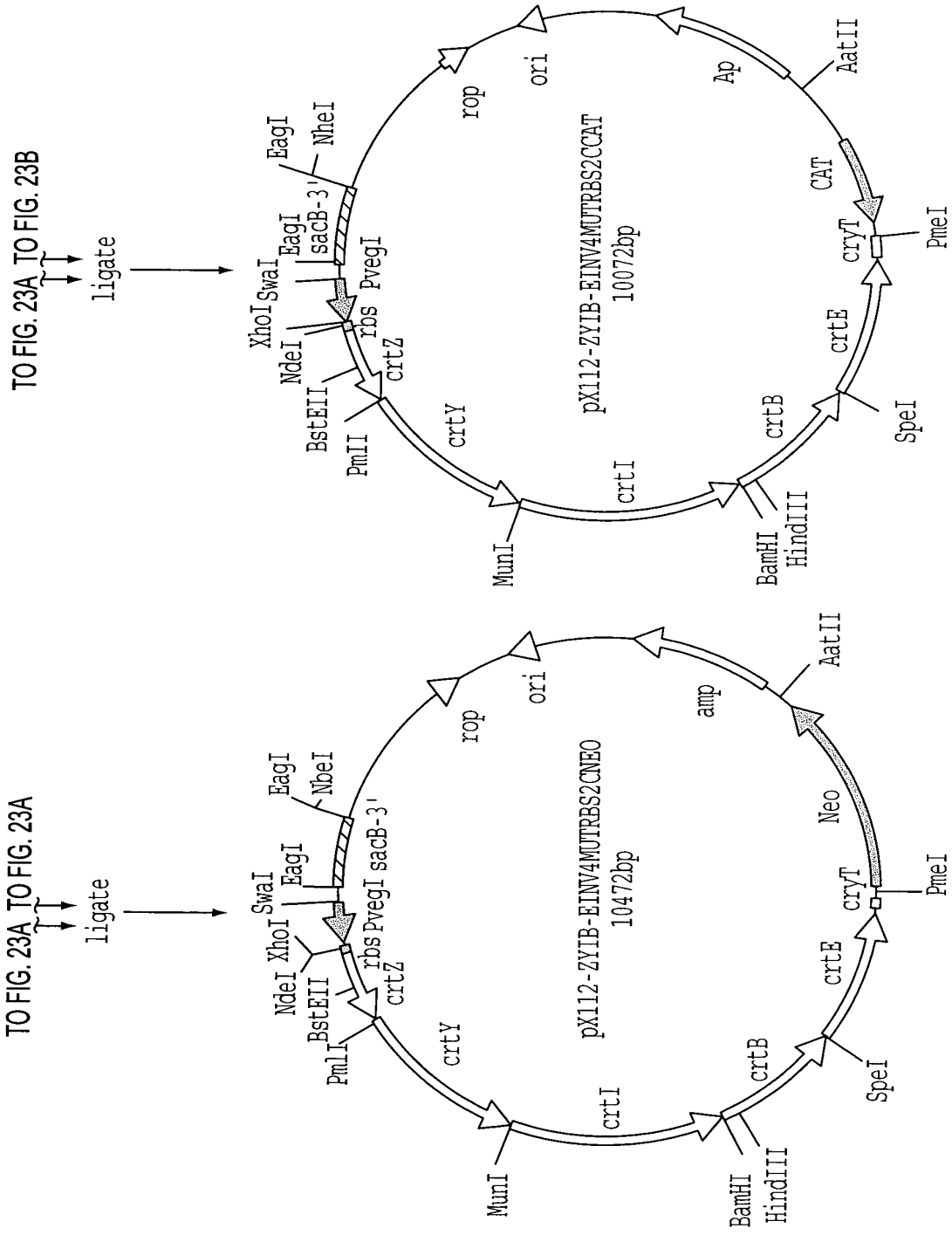

TO FIG. 23A TO FIG. 23B
ligate
pX112-ZYIB-EINV4MUTRBS2CCAT
10072bp
rop
ori
Ap
AatII
CAT
cryT
PmeI
crtE
SpeI
crtB
HindIII
BamHI
crtI
MunI
crtY
crtZ
rbs
PvegI
sacB-3'
EagI
NheI
XhoI
SwaI
NdeI
BstEII
PmlI
TO FIG. 23A TO FIG. 23A
ligate
pX112-ZYIB-EINV4MUTRBS2CNEO
10472bp
rop
ori
amp
AatII
Neo
cryT
PmeI
crtE
SpeI
crtB
HindIII
BamHI
crtI
MunI
crtY
crtZ
rbs
PvegI
sacB-3'
EagI
NbeI
SwaI
XhoI
NdeI
BstEII
PmlI

FIG. 23C

```
   1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC  60
     GATTTAACATTCGCAATTATAAAACAATTTTAAGCGCAATTTAAAAACAATTTAGTCGAG

  61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
     TAAAAAATTGGTTATCCGGCTTTAGCCGTTTTAGGGAATATTTAGTTTTCTTATCTGGCT

 121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
     CTATCCCAACTCACAACAAGGTCAAACCTTGTTCTCAGGTGATAATTTCTTGCACCTGAG

 181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
     GTTGCAGTTTCCCGCTTTTTGGCAGATAGTCCCGCTACCGGGTGATGCACTTGGTAGTGG

 241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
     GATTAGTTCAAAAAACCCCAGCTCCACGGCATTTCGTGATTTAGCCTTGGGATTTCCCTC

 301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
     GGGGGCTAAATCTCGAACTGCCCCTTTCGGCCGCTTGCACCGCTCTTTCCTTCCCTTCTT

 361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
     TCGCTTTCCTCGCCCGCGATCCCGCGACCGTTCACATCGCCAGTGCGACGCGCATTGGTG

 421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
     GTGTGGGCGGCGCGAATTACGCGGCGATGTCCCGCGCAGGGTAAGCGGTAAGTCCGACGC

 481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
     GTTGACAACCCTTCCCGCTAGCCACGCCCGGAGAAGCGATAATGCGGTCGACCGCTTTCC

 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
     CCCTACACGACGTTCCGCTAATTCAACCCATTGCGGTCCCAAAAGGGTCAGTGCTGCAAC

 601 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 660
     ATTTTGCTGCCGGTCACTCGCGCGCATTATGCTGAGTGATATCCCGCTTAACCTCGAGGT

 661 CCGCGGTGGCGGCCGCTCTAGTGGATCCGCGCCTGGCCGTTCGCGATCAGCAGCCGCCCT 720
     GGCGCCACCGCCGGCGAGATCACCTAGGCGCGGACCGGCAAGCGCTAGTCGTCGGCGGGA

 721 TGCGGATCGGTCAGCATCATCCCCATGAACCGCAGCGCACGACGCAGCGCGCGCCCCAGA 780
     ACGCCTAGCCAGTCGTAGTAGGGGTACTTGGCGTCGCGTGCTGCGTCGCGCGCGGGGTCT

 781 TCGGGCGCGTCCAGCACGGCATGCGCCATCATCGCGAAGGCCCCCGGCGGCATGGGGCGC 840
     AGCCCGCGCAGGTCGTGCCGTACGCGGTAGTAGCGCTTCCGGGGGCCGCCGTACCCCGCG

 841 GTGCCCATTCCGAAGAACTCGCAGCCTGTCCGCTGCGCAAGGTCGCGCCAGATCGCGCCG 900
     CACGGGTAAGGCTTCTTGAGCGTCGGACAGGCGACGCGTTCCAGCGCGGTCTAGCGCGGC

 901 TATTCCGATGCAGTGACGGGCCCGATGCGCGTGGGCCCGCCCTGCCCCGCCGCCACCAGC 960
     ATAAGGCTACGTCACTGCCCGGGCTACGCGCACCCGGGCGGGACGGGGCGGCGGTGGTCG
```

FIG. 24A

```
961  GCATCGCGCACGAACCCTTCCGAGATGATGTGCTGATCCATGGCCCGTCATTGCAAAACC 1020
     CGTAGCGCGTGCTTGGGAAGGCTCTACTACACGACTAGGTACCGGGCAGTAACGTTTTGG

1021 GATCACCGATCCTGTCGCGTGATGGCATTGTTTGCAATGCCCCGAGGGCTAGGATGGCGC 1080
     CTAGTGGCTAGGACAGCGCACTACCGTAACAAACGTTACGGGGCTCCCGATCCTACCGCG

1081 GAAGGATCAAGGGGGGGAGAGACATGGAAATCGAGGGACGGGTCTTTGTCGTCACGGGCG 1140
     CTTCCTAGTTCCCCCCCTCTCTGTACCTTTAGCTCCCTGCCCAGAAACAGCAGTGCCCGC

1141 CCGCATCGGGTCTGGGGGCGGCCTCGGCGCGGATGCTGGCCCAAGGCGGCGCGAAGGTCG 1200
     GGCGTAGCCCAGACCCCCGCCGGAGCCGCGCCTACGACCGGGTTCCGCCGCGCTTCCAGC

1201 TGCTGGCCGATCTGGCGGAACCGAAGGACGCGCCCGAAGGCGCGGTTCACGCGGCCTGCG 1260
     ACGACCGGCTAGACCGCCTTGGCTTCCTGCGCGGGCTTCCGCGCCAAGTGCGCCGGACGC

1261 ACGTGACCGACGCGACCGCTGCGCAGACGGCCATCGCGCTGGCGACCGACCGCTTCGGCA 1320
     TGCACTGGCTGCGCTGGCGACGCGTCTGCCGGTAGCGCGACCGCTGGCTGGCGAAGCCGT

1321 GGCTGGACGGCCTTGTGAACTGCGCGGGCATCGCGCCGGCCGAACGGATGCTGGGCCGCG 1380
     CCGACCTGCCGGAACACTTGACGCGCCCGTAGCGCGGCCGGCTTGCCTACGACCCGGCGC

1381 ACGGGCCGCATGGACTGGACAGCTTTGCCCGTGCGGTCACGATCAACCTGATCGGCAGCT 1440
     TGCCCGGCGTACCTGACCTGTCGAAACGGGCACGCCAGTGCTAGTTGGACTAGCCGTCGA

1441 TCAACATGGCCCGCCTTGCAGCCGAGGCGATGGCCCGGAACGAGCCCGTCCGGGGCGAGC 1500
     AGTTGTACCGGGCGGAACGTCGGCTCCGCTACCGGGCCTTGCTCGGGCAGGCCCCGCTCG

1501 GTGGCGTGATCGTCAACACGGCCTCGATCGCGGCGCAGGACGGACAGATCGGACAGGTCG 1560
     CACCGCACTAGCAGTTGTGCCGGAGCTAGCGCCGCGTCCTGCCTGTCTAGCCTGTCCAGC

1561 CCTATGCGGCCAGCAAGGCGGGCGTGGCGGGCATGACGCTGCCGATGGCCCGCGACCTTG 1620
     GGATACGCCGGTCGTTCCGCCCGCACCGCCCGTACTGCGACGGCTACCGGGCGCTGGAAC

1621 CGCGGCACGGCATCCGCGTCATGACCATCGCGCCCGGCATCTTCCGCACCCCGATGCTGG 1680
     GCGCCGTGCCGTAGGCGCAGTACTGGTAGCGCGGGCCGTAGAAGGCGTGGGGCTACGACC

1681 AGGGGCTGCCGCAGGACGTTCAGGACAGCCTGGGCGCGGCGGTGCCCTTCCCCTCGCGGC 1740
     TCCCCGACGGCGTCCTGCAAGTCCTGTCGGACCCGCGCCGCCACGGGAAGGGGAGCGCCG

1741 TGGGAGAGCCGTCGGAATACGCGGCGCTGTTGCACCACATCATCGCGAACCCCATGCTGA 1800
     ACCCTCTCGGCAGCCTTATGCGCCGCGACAACGTGGTGTAGTAGCGCTTGGGGTACGACT

1801 ACGGAGAGGTCATCCGCCTCGACGGCGCATTGCGCATGGCCCCCAAGTGAAGGAGCGTTT 1860
     TGCCTCTCCAGTAGGCGGAGCTGCCGCGTAACGCGTACCGGGGGTTCACTTCCTCGCAAA

1861 CATGGACCCCATCGTCATCACCGGCGCGATGCGCACCCCGATGGGGGCATTCCAGGGCGA 1920
     GTACCTGGGGTAGCAGTAGTGGCCGCGCTACGCGTGGGGCTACCCCCGTAAGGTCCCGCT

1921 TCTTGCCGCGATGGATGCCCCGACCCTTGGCGCGGACGCGATCCGCGCCGCGCTGAACGG 1980
     AGAACGGCGCTACCTACGGGGCTGGGAACCGCGCCTGCGCTAGGCGCGGCGCGACTTGCC
```

FIG. 24B

```
1981 CCTGTCGCCCGACATGGTGGACGAGGTGCTGATGGGCTGCGTCCTCGCCGCGGGCCAGGG 2040
     GGACAGCGGGCTGTACCACCTGCTCCACGACTACCCGACGCAGGAGCGGCGCCCGGTCCC

2041 TCAGGCACCGGCACGTCAGGCGGCGCTTGGCGCCGGACTGCCGCTGTCGACGGGCACGAC 2100
     AGTCCGTGGCCGTGCAGTCCGCCGCGAACCGCGGCCTGACGGCGACAGCTGCCCGTGCTG

2101 CACCATCAACGAGATGTGCGGATCGGGCATGAAGGCCGCGATGCTGGGCCATGACCTGAT 2160
     GTGGTAGTTGCTCTACACGCCTAGCCCGTACTTCCGGCGCTACGACCCGGTACTGGACTA

2161 CGCCGCGGGATCGGCGGGCATCGTCGTCGCCGGCGGGATGGAGAGCATGTCGAACGCCCC 2220
     GCGGCGCCCTAGCCGCCCGTAGCAGCAGCGGCCGCCCTACCTCTCGTACAGCTTGCGGGG

2221 CTACCTGCTGCCCAAGGCGCGGTCGGGGATGCGCATGGGCCATGACCGTGTGCTGGATCA 2280
     GATGGACGACGGGTTCCGCGCCAGCCCCTACGCGTACCCGGTACTGGCACACGACCTAGT

2281 CATGTTCCTCGACGGGTTGGAGGACGCCTATGACAAGGGCCGCCTGATGGGCACCTTCGC 2340
     GTACAAGGAGCTGCCCAACCTCCTGCGGATACTGTTCCCGGCGGACTACCCGTGGAAGCG

2341 CGAGGATTGCGCCGGCGATCACGGTTTCACCCGCGAGGCGCAGGACGACTATGCGCTGAC 2400
     GCTCCTAACGCGGCCGCTAGTGCCAAAGTGGGCGCTCCGCGTCCTGCTGATACGCGACTG

2401 CAGCCTGGCCCGCGCGCAGGACGCCATCGCCAGCGGTGCCTTCGCCGCCGAGATCGCGCC 2460
     GTCGGACCGGGCGCGCGTCCTGCGGTAGCGGTCGCCACGGAAGCGGCGGCTCTAGCGCGG

2461 CGTGACCGTCACGGCACGCAAGGTGCAGACCACCGTCGATACCGACGAGATGCCCGGCAA 2520
     GCACTGGCAGTGCCGTGCGTTCCACGTCTGGTGGCAGCTATGGCTGCTCTACGGGCCGTT

2521 GGCCCGCCCCGAGAAGATCCCCCATCTGAAGCCCGCCTTCCGTGACGGTGGCACGGTCAC 2580
     CCGGGCGGGGCTCTTCTAGGGGGTAGACTTCGGGCGGAAGGCACTGCCACCGTGCCAGTG

2581 GGCGGCGAACAGCTCGTCGATCTCGGACGGGGCGGCGGCGCTGGTGATGATGCGCCAGTC 2640
     CCGCCGCTTGTCGAGCAGCTAGAGCCTGCCCCGCCGCCGCGACCACTACTACGCGGTCAG

2641 GCAGGCCGAGAAGCTGGGCCTGACGCCGATCGCGCGGATCATCGGTCATGCGACCCATGC 2700
     CGTCCGGCTCTTCGACCCGGACTGCGGCTAGCGCGCCTAGTAGCCAGTACGCTGGGTACG

2701 CGACCGTCCCGGCCTGTTCCCGACGGCCCCCATCGGCGCGATGCGCAAGCTGCTGGACCG 2760
     GCTGGCAGGGCCGGACAAGGGCTGCCGGGGGTAGCCGCGCTACGCGTTCGACGACCTGGC

2761 CACGGACACCCGCCTTGGCGATTACGACCTGTTCGAGGTGAACGAGGCATTCGCCGTCGT 2820
     GTGCCTGTGGGCGGAACCGCTAATGCTGGACAAGCTCCACTTGCTCCGTAAGCGGCAGCA

2821 CGCCATGATCGCGATGAAGGAGCTTGGCCTGCCACACGATGCCACGAACATCAACGGCGG 2880
     GCGGTACTAGCGCTACTTCCTCGAACCGGACGGTGTGCTACGGTGCTTGTAGTTGCCGCC

2881 GGCCTGCGCGCTTGGGCATCCCATCGGCGCGTCGGGGGCGCGGATCATGGTCACGCTGCT 2940
     CCGGACGCGCGAACCCGTAGGGTAGCCGCGCAGCCCCCGCGCCTAGTACCAGTGCGACGA

2941 GAACGCGATGGCGGCGCGGGGCGCGACGCGCGGGGCCGCATCCGTCTGCATCGGCGGGGG 3000
     CTTGCGCTACCGCCGCGCCCCGCGCTGCGCGCCCCGGCGTAGGCAGACGTAGCCGCCCCC
```

FIG. 24C

```
3001 CGAGGCGACGGCCATCGCGCTGGAACGGCTGAGCTAATTCATTTGCGCGAATCCGCGTTT 3060
     GCTCCGCTGCCGGTAGCGCGACCTTGCCGACTCGATTAAGTAAACGCGCTTAGGCGCAAA

3061 TTCGTGCACGATGGGGGAACCGGAAACGGCCACGCCTGTTGTGGTTGCGTCGACCTGTCT 3120
     AAGCACGTGCTACCCCCTTGGCCTTTGCCGGTGCGGACAACACCAACGCAGCTGGACAGA

3121 TCGGGCCATGCCCGTGACGCGATGTGGCAGGCGCATGGGGCGTTGCCGATCCGGTCGCAT 3180
     AGCCCGGTACGGGCACTGCGCTACACCGTCCGCGTACCCCGCAACGGCTAGGCCAGCGTA

3181 GACTGACGCAACGAAGGCACCGATGACGCCCAAGCAGCAATTCCCCCTACGCGATCTGGT 3240
     CTGACTGCGTTGCTTCCGTGGCTACTGCGGGTTCGTCGTTAAGGGGGATGCGCTAGACCA

3241 CGAGATCAGGCTGGCGCAGATCTCGGGCCAGTTCGGCGTGGTCTCGGCCCCGCTCGGCGC 3300
     GCTCTAGTCCGACCGCGTCTAGAGCCCGGTCAAGCCGCACCAGAGCCGGGGCGAGCCGCG

3301 GGCCATGAGCGATGCCGCCCTGTCCCCCGGCAAACGCTTTCGCGCCGTGCTGATGCTGAT 3360
     CCGGTACTCGCTACGGCGGGACAGGGGGCCGTTTGCGAAAGCGCGGCACGACTACGACTA

3361 GGTCGCCGAAAGCTCGGGCGGGGTCTGCGATGCGATGGTCGATGCCGCCTGCGCGGTCGA 3420
     CCAGCGGCTTTCGAGCCCGCCCCAGACGCTACGCTACCAGCTACGGCGGACGCGCCAGCT

3421 GATGGTCCATGCCGCATCGCTGATCTTCGACGACATGCCCTGCATGGACGATGCCAGGAC 3480
     CTACCAGGTACGGCGTAGCGACTAGAAGCTGCTGTACGGGACGTACCTGCTACGGTCCTG

3481 CCGTCGCGGTCAGCCCGCCACCCATGTCGCCCATGGCGAGGGGCGCGCGGTGCTTGCGGG 3540
     GGCAGCGCCAGTCGGGCGGTGGGTACAGCGGGTACCGCTCCCCGCGCGCCACGAACGCCC

3541 CATCGCCCTGATCACCGAGGCCATGCGGATTTTGGGCGAGGCGCGCGGCGCGACGCCGGA 3600
     GTAGCGGGACTAGTGGCTCCGGTACGCCTAAAACCCGCTCCGCGCGCCGCGCTGCGGCCT

3601 TCAGCGCGCAAGGCTGGTCGCATCCATGTCGCGCGCGATGGGACCGGTGGGGCTGTGCGC 3660
     AGTCGCGCGTTCCGACCAGCGTAGGTACAGCGCGCGCTACCCTGGCCACCCCGACACGCG

3661 AGGGCAGGATCTGGACCTGCACGCCCCCAAGGACGCCGCCGGGATCGAACGTGAACAGGA 3720
     TCCCGTCCTAGACCTGGACGTGCGGGGGTTCCTGCGGCGGCCCTAGCTTGCACTTGTCCT

3721 CCTCAAGACCGGCGTGCTGTTCGTCGCGGGCCTCGAGATGCTGTCCATTATTAAGGGTCT 3780
     GGAGTTCTGGCCGCACGACAAGCAGCGCCCGGAGCTCTACGACAGGTAATAATTCCCAGA

3781 GGACAAGGCCGAGACCGAGCAGCTCATGGCCTTCGGGCGTCAGCTTGGTCGGGTCTTCCA 3840
     CCTGTTCCGGCTCTGGCTCGTCGAGTACCGGAAGCCCGCAGTCGAACCAGCCCAGAAGGT

3841 GTCCTATGACGACCTGCTGGACGTGATCGGCGACAAGGCCAGCACCGGCAAGGATACGGC 3900
     CAGGATACTGCTGGACGACCTGCACTAGCCGCTGTTCCGGTCGTGGCCGTTCCTATGCCG

3901 GCGCGACACCGCCGCCCCCGGCCCAAAGGGCGGCCTGATGGCGGTCGGACAGATGGGCGA 3960
     CGCGCTGTGGCGGCGGGGGCCGGGTTTCCCGCCGGACTACCGCCAGCCTGTCTACCCGCT

3961 CGTGGCGCAGCATTACCGCGCCAGCCGCGCGCAACTGGACGAGCTGATGCGCACCCGGCT 4020
     GCACCGCGTCGTAATGGCGCGGTCGGCGCGCGTTGACCTGCTCGACTACGCGTGGGCCGA
```

FIG. 24D

```
4021 GTTCCGCGGGGGGCAGATCGCGGACCTGCTGGCCCGCGTGCTGCCGCATGACATCCGCCG 4080
     CAAGGCGCCCCCCGTCTAGCGCCTGGACGACCGGGCGCACGACGGCGTACTGTAGGCGGC

4081 CAGCGCCTAGGCGCGCGGTCGGGTCCACAGGCCGTCGCGGCTGATTTCGCCGCCGCGCAG 4140
     GTCGCGGATCCGCGCGCCAGCCCAGGTGTCCGGCAGCGCCGACTAAAGCGGCGGCGCGTC

4141 GCGCGATGCGGCCGCGTCCAAGCCTCCGCGCGCCAGAAGCCCGATCTTGGCAGCCTTCGA 4200
     CGCGCTACGCCGGCGCAGGTTCGGAGGCGCGCGGTCTTCGGGCTAGAACCGTCGGAAGCT

4201 CGTGCTGATCCGCTGGCGATAGGCCTCGGGGCCACCCTGCCGGATGCGCGTCCCGATTGC 4260
     GCACGACTAGGCGACCGCTATCCGGAGCCCCGGTGGGACGGCCTACGCGCAGGGCTAACG

4261 GCGATAGATACGCAGCGCGGCGGCGATCGACCACGCGCAGCGCGGCGGCAGATGCGGAAG 4320
     CGCTATCTATGCGTCGCGCCGCCGCTAGCTGGTGCGCGTCGCGCCGCCGTCTACGCCTTC

4321 CCCCTGCCGCGCCGAGGCATAATAGGGCTCGGCCGCGTCAAGCAGGCGGATGATGACGGA 4380
     GGGGACGGCGCGGCTCCGTATTATCCCGAGCCGGCGCAGTTCGTCCGCCTACTACTGCCT

4381 ATAGAGCGCGTCCGAAGGCACCGGACCCTCAACCGTCGCCCCCGCCTCGGCCAGCCAGTC 4440
     TATCTCGCGCAGGCTTCCGTGGCCTGGGAGTTGGCAGCGGGGGCGGAGCCGGTCGGTCAG

4441 GGCAGGCAGATAGCAGCGCCCGATGGCGGCATCGTCGATCACGTCGCGAGCGATGTTCGT 4500
     CCGTCCGTCTATCGTCGCGGGCTACCGCCGTAGCAGCTAGTGCAGCGCTCGCTACAAGCA

4501 CAGCTGGAACGCAAGGCCCAGATCGCAGGCGCGATCCAGCACCGCATCGTCCTGCACGCC 4560
     GTCGACCTTGCGTTCCGGGTCTAGCGTCCGCGCTAGGTCGTGGCGTAGCAGGACGTGCGG

4561 CATCACCCGCGCCATCATCACGCCCACGACCCCCGCGACGTGGTAGGAATATTCCAGCAC 4620
     GTAGTGGGCGCGGTAGTAGTGCGGGTGCTGGGGGCGCTGCACCATCCTTATAAGGTCGTG

4621 GTCATCCAGGCTGCGGTATTCGCGATCCGCGACATCCATCGCGAAACCCTCGATCAGGTC 4680
     CAGTAGGTCCGACGCCATAAGCGCTAGGCGCTGTAGGTAGCGCTTTGGGAGCTAGTCCAG

4681 CATCGGCCAAAGGTCCGGGAAATCATGCCGCCGGGCGACCTGGCGCAGCGCCGCGAAGGG 4740
     GTAGCCGGTTTCCAGGCCCTTTAGTACGGCGGCCCGCTGGACCGCGTCGCGGCGCTTCCC

4741 CGGCGACATCGGGCCGTCCTCGTGCAGCGCGGCCAGCGTGTCGGCGCGCAGCGCCCCCAG 4800
     GCCGCTGTAGCCCGGCAGGAGCACGTCGCGCCGGTCGCACAGCCGCGCGTCGCGGGGGTC

4801 CCGCGCCTGTGGGTCGCCGCCCGCCTCGGGGGCAGAACCCATCACCTGCCCGTCGATCAC 4860
     GGCGCGGACACCCAGCGGCGGGCGGAGCCCCCGTCTTGGGTAGTGGACGGGCAGCTAGTG

4861 GTCATCCGCATGCCTGCACCAGGCATAGAGCATGACCGTATCCTCGCGGATGCCGGGCGG 4920
     CAGTAGGCGTACGGACGTGGTCCGTATCTCGTACTGGCATAGGAGCGCCTACGGCCCGCC

4921 CATCAGCTTGGCCGCCTGCGCGAAGCTTTGCGAACCCTGCGCGATGGCCGCTTCGGAAGT 4980
     GTAGTCGAACCGGCGGACGCGCTTCGAAACGCTTGGGACGCGCTACCGGCGAAGCCTTCA

4981 CGCCGTCAGATCGGTCATGCGACGGCCAGGTCCGACAGCATGACCTGCGCCGTGGCCTTG 5040
     GCGGCAGTCTAGCCAGTACGCTGCCGGTCCAGGCTGTCGTACTGGACGCGGCACCGGAAC
```

FIG. 24E

```
5041 GCGCTGCCAACGACACCCGGGATGCCCGCACCCGGATGCGTGCCCGCCCCCACGATGTAG 5100
     CGCGACGGTTGCTGTGGGCCCTACGGGCGTGGGCCTACGCACGGGCGGGGGTGCTACATC

5101 AAGTTCGGGATCGCGCGGTCGCGGTTATGCGGGCGGAACCAGGCGGATTGCGTCAGGATC 5160
     TTCAAGCCCTAGCGCGCCAGCGCCAATACGCCCGCCTTGGTCCGCCTAACGCAGTCCTAG

5161 GGCTCGACCGAGAAGGCGCTGCCGTGATGGGCCGACAGTTCGGTGCTGAAATCGGCGGGG 5220
     CCGAGCTGGCTCTTCCGCGACGGCACTACCCGGCTGTCAAGCCACGACTTTAGCCGCCCC

5221 CTGAAGATGCGGCTGACGGTCAGGTGCTTGCGCAGGTCGGGGATGGCGCGGCGCTCCAGT 5280
     GACTTCTACGCCGACTGCCAGTCCACGAACGCGTCCAGCCCCTACCGCGCCGCGAGGTCA

5281 TCCTCGAAGATGCGCTCGGCATAGCCCGGGGCCTCGGCTTCCCAATCGACATCGGCGCGG 5340
     AGGAGCTTCTACGCGAGCCGTATCGGGCCCCGGAGCCGAAGGGTTAGCTGTAGCCGCGCC

5341 CCCAGATGCGGAACGGGCGCAAGGACGTAATGCGTGGACATCCCCTCGGGGGCCAGGCTG 5400
     GGGTCTACGCCTTGCCCGCGTTCCTGCATTACGCACCTGTAGGGGAGCCCCCGGTCCGAC

5401 GGATCGGTCACGCAGGGCGAATGCAGATACATCGAGAAATCGTCCGGCAGGCGTGGCCCG 5460
     CCTAGCCAGTGCGTCCCGCTTACGTCTATGTAGCTCTTTAGCAGGCCGTCCGCACCGGGC

5461 TTGAAGATCTCGTTCACCAGCCCCTTGTAGCGCGGGCCGAAGATGACGCTGTGGTGGGCC 5520
     AACTTCTAGAGCAAGTGGTCGGGGAACATCGCGCCCGGCTTCTACTGCGACACCACCCGG

5521 AGGTTCTCGGGGCGCTTGGACAGGCCGAAATGCAGCACGAACAGCGACATCGACCAGCGC 5580
     TCCAAGAGCCCCGCGAACCTGTCCGGCTTTACGTCGTGCTTGTCGCTGTAGCTGGTCGCG

5581 TGCCGGTTCAGGATCGCGGCCTTGGTGCGCCCGCGGCGGGTATGGCCCAGCAGGTCGCGA 5640
     ACGGCCAAGTCCTAGCGCCGGAACCACGCGGGCGCCGCCCATACCGGGTCGTCCAGCGCT

5641 TAGCTGTGCATCACGTCGCCGTTGCTGGCCACCGTATCCGCGCGCAACTGCCGCCCGTCC 5700
     ATCGACACGTAGTGCAGCGGCAACGACCGGTGGCATAGGCGCGCGTTGACGGCGGGCAGG

5701 AGCAGCGTGACGCCCGTGGCGCGATCGCCCTCGGTGTCGATCCGCGTGACGCGGGCATTC 5760
     TCGTCGCACTGCGGGCACCGCGCTAGCGGGAGCCACAGCTAGGCGCACTGCGCCCGTAAG

5761 AGCAGCAGCGTGCCGCCAAGACGCTCGAACAGGGCGACCATGCCCGCGACCAGCTGGTTG 5820
     TCGTCGTCGCACGGCGGTTCTGCGAGCTTGTCCCGCTGGTACGGGCGCTGGTCGACCAAC

5821 GTGCCGCCCTTGGCGAACCAGACGCCGCCGCGCCGTTCCAGCGCATGGATCAGCGCATAG 5880
     CACGGCGGGAACCGCTTGGTCTGCGGCGGCGCGGCAAGGTCGCGTACCTAGTCGCGTATC

5881 ATCGAGCTGGTCGAAAACGGGTTCCCGCCGACCAGCAGCGTGTGGAACGAGAAGGCCTGC 5940
     TAGCTCGACCAGCTTTTGCCCAAGGGCGGCTGGTCGTCGCACACCTTGCTCTTCCGGACG

5941 CGCAGATGCGGGTCCTGGATGAAGCGCGCCACCATGCTGTGGACCGAGCGGTATGCCTGC 6000
     GCGTCTACGCCCAGGACCTACTTCGCGCGGTGGTACGACACCTGGCTCGCCATACGGACG

6001 AGGCGCATCAGCGCCGGCGCGGCGTTCAGCATCTGGCCCAGCTTCAGGAAGGGCGTGGTC 6060
     TCCGCGTAGTCGCGGCCGCGCCGCAAGTCGTAGACCGGGTCGAAGTCCTTCCCGCACCAG
```

FIG. 24F

```
6061 CCCAGCTTCAGATACCCCTCGCGATAGACCTCCTCGGCGTAATCGTGGAAGCGGCGATAG 6120
     GGGTCGAAGTCTATGGGGAGCGCTATCTGGAGGAGCCGCATTAGCACCTTCGCCGCTATC

6121 CCATCGACATCGGCGGGATTGAAGGAGGCGACCTGGCGGATCAGCTCGTCGTCGTCGTTC 6180
     GGTAGCTGTAGCCGCCCTAACTTCCTCCGCTGGACCGCCTAGTCGAGCAGCAGCAGCAAG

6181 ACGTATTCGAAGCTGCGGCCGTCCGCCCATGTCAGCCGGTAGAAGGGCGAGACCGGCAGC 6240
     TGCATAAGCTTCGACGCCGGCAGGCGGGTACAGTCGGCCATCTTCCCGCTCTGGCCGTCG

6241 AGCGTCACGTCACGCTCCATCGGTTGGCCGCTGAGGGCCCACAGCTCTCGCAGGCTGTCG 6300
     TCGCAGTGCAGTGCGAGGTAGCCAACCGGCGACTCCCGGGTGTCGAGAGCGTCCGACAGC

6301 GGGTCGGTCACGACCGTCGGGCCTGCATCGAAGACGTGGCCCTGATCGTTCCAGACATAG 6360
     CCCAGCCAGTGCTGGCAGCCCGGACGTAGCTTCTGCACCGGGACTAGCAAGGTCTGTATC

6361 GCGCGGCCGCCGGGCTTGTCGCGGGCCTCGACGATGGTGGTCGCGATGCCGGCCGATTGC 6420
     CGCGCCGGCGGCCCGAACAGCGCCCGGAGCTGCTACCACCAGCGCTACGGCCGGCTAACG

6421 AGGCGGATGGCAAGCGCAAGCCCGCCGAAACCTGCGCCGATGACGATGGCGGAACTCATG 6480
     TCCGCCTACCGTTCGCGTTCGGGCGGCTTTGGACGCGGCTACTGCTACCGCCTTGAGTAC

6481 CTCTCTCCTGCAGCAGGGGGCGTTCGGGCAGGCAGCGCACGGCCTGCGACAGCGGAATGG 6540
     GAGAGAGGACGTCGTCCCCCGCAAGCCCGTCCGTCGCGTGCCGGACGCTGTCGCCTTACC

6541 GCGGGCGTCCGGTGACGATGCGAAGCCGGTCGGCCAATGTCAGGCGCCCGGCATAGAAGC 6600
     CGCCCGCAGGCCACTGCTACGCTTCGGCCAGCCGGTTACAGTCCGCGGGCCGTATCTTCG

6601 GCTCGATCAGCGGCTGCGGCAGGCGGTAGAACCGCTGCAGCAGGCGATAGCGACGGTCGG 6660
     CGAGCTAGTCGCCGACGCCGTCCGCCATCTTGGCGACGTCGTCCGCTATCGCTGCCAGCC

6661 GCGGGCAGCCGCGGAACAGCATCCGGTTCAGCAGCCGCAGGAAGCGGTCGCGATCCGCGC 6720
     CGCCCGTCGGCGCCTTGTCGTAGGCCAAGTCGTCGGCGTCCTTCGCCAGCGCTAGGCGCG

6721 GATCGATGGCCCAGCCGCGCACCGCGCGACGGGCGGACGCGGTCGTCAGGTCGCGCGCCG 6780
     CTAGCTACCGGGTCGGCGCGTGGCGCGCTGCCCGCCTGCGCCAGCAGTCCAGCGCGCGGC

6781 CGATGGCATCCGCGACCTGCGCGGCATAGGGCAGCGAATATCCGGTGACGGGGTGGAACA 6840
     GCTACCGTAGGCGCTGGACGCGCCGTATCCCGTCGCTTATAGGCCACTGCCCCACCTTGT

6841 GCCCTGCCCCCAGCCCAACCGGCACCGCCCCCTGCGCGTGGTCGCGCCAGAAGCCTATGG 6900
     CGGGACGGGGGTCGGGTTGGCCGTGGCGGGGGACGCGCACCAGCGCGGTCTTCGGATACC

6901 CGTCATGGGCCAGCGCGATGGGCAGGATGCCCCTTTCGCGCCGCATCTCCTGCCCGGTCC 6960
     GCAGTACCCGGTCGCGCTACCCGTCCTACGGGGAAAGCGCGGCGTAGAGGACGGGCCAGG

6961 AGCCCCGCCTGGCGGCATAGTCCAGCGACGCCTGCGCCAGCGCGCCATCGTCCAGATCGC 7020
     TCGGGGCGGACCGCCGTATCAGGTCGCTGCGGACGCGGTCGCGCGGTAGCAGGTCTAGCG
```

FIG. 24G

```
7021 CGCCGTCGCTGTAGCGCGTATCCTCGATCAGGATGCGGGTGGGACTGAAGGGCAGCAGAT 7080
     GCGGCAGCGACATCGCGCATAGGAGCTAGTCCTACGCCCACCCTGACTTCCCGTCGTCTA

7081 AGATGAAGCGGTACCCGTCCATCTGCGGAACGGTCGCGTCCATGATCATCGGGCGCTCGA 7140
     TCTACTTCGCCATGGGCAGGTAGACGCCTTGCCAGCGCAGGTACTAGTAGCCCGCGAGCT

7141 CGCCATGGGGGGCGTCGGTCTCGATCTCGACGCCCACGAATTTCTGGAAACCCACGGTCA 7200
     GCGGTACCCCCCGCAGCCAGAGCTAGAGCTGCGGGTGCTTAAAGACCTTTGGGTGCCAGT

7201 GGTGCGGGGTCTCGACGGCACCACGGGCGTCGATCACGCAGGCAGCCTCGATCCGCGAGC 7260
     CCACGCCCCAGAGCTGCCGTGGTGCCCGCAGCTAGTGCGTCCGTCGGAGCTAGGCGCTCG

7261 CGTCCGTCAGCGTCGCGCCGGTATCGTCCAGCGTCGCGACATGCGTATTCCACCGCAGAT 7320
     GCAGGCAGTCGCAGCGCGGCCATAGCAGGTCGCAGCGCTGTACGCATAAGGTGGCGTCTA

7321 CGACACCCTGCAGCAGCCCGATCAGCGCGCCCGCCTCGATCGAGCCATAGCCTGTCGTCA 7380
     GCTGTGGGACGTCGTCGGGCTAGTCGCGCGGGCGGAGCTAGCTCGGTATCGGACAGCAGT

7381 GGCGGCGCGAATGGTCGGGAAACGCGACCTCCTGATCCGTCCATTCGCCGCGACGAATGG 7440
     CCGCCGCGCTTACCAGCCCTTTGCGCTGGAGGACTAGGCAGGTAAGCGGCGCTGCTTACC

7441 GCGACAGGCGCGCCAGCCATTCGGGCGAAAGATCCGTGTCGTGGCAGGACCAGGTGTGCT 7500
     CGCTGTCCGCGCGGTCGGTAAGCCCGCTTTCTAGGCACAGCACCGTCCTGGTCCACACGA

7501 GGTCCGAGGGGCCGGACCGCGCGTCGAGCATCACGATGCGCGCATCCGGTCTGCGGTCGC 7560
     CCAGGCTCCCCGGCCTGGCGCGCAGCTCGTAGTGCTACGCGCGTAGGCCAGACGCCAGCG

7561 GAACGGCAAGCGCGATCAGCGCACCGGACAGCCCCGCGCCCGCGATCAGCAGATCATGGC 7620
     CTTGCCGTTCGCGCTAGTCGCGTGGCCTGTCGGGGCGCGGGCGCTAGTCGTCTAGTACCG

7621 TCATGTATTGCGATCCGCCCCTTCGCGGTCCTTCAGCAGCGCGCCCGAGCGTTTCAGCTC 7680
     AGTACATAACGCTAGGCGGGGAAGCGCCAGGAAGTCGTCGCGCGGGCTCGCAAAGTCGAG

7681 TGCCTTGAGGCTGTCGACCGAGGGCGCCCAGATGAAACCGAAGCTGACGCAGTTCTCGCG 7740
     ACGGAACTCCGACAGCTGGCTCCCGCGGGTCTACTTTGGCTTCGACTGCGTCAAGAGCGC

7741 GCCATGGACCGCGTGATGCATCCTGTGTGCCTGGTAGACGCGACGAAGATAGCCGCGCTT 7800
     CGGTACCTGGCGCACTACGTAGGACACACGGACCATCTGCGCTGCTTCTATCGGCGCGAA

7801 GGGGACATAGCGGAACGGCCAGCGCCCATGCACCAAGCCGTCATGCAGGAAATAGTAGAT 7860
     CCCCTGTATCGCCTTGCCGGTCGCGGGTACGTGGTTCGGCAGTACGTCCTTTATCATCTA

7861 CAGCCCGTAGCAGGTGACCCCCACCGCCAGCCACCAGGCCAGATCCGACCCCATCGCGCC 7920
     GTCGGGCATCGTCCACTGGGGGTGGCGGTCGGTGGTCCGGTCTAGGCTGGGGTAGCGCGG

7921 GATCGCGAACAGCACGATCGAGATTACCGCGAAGATGACGCCATAGAGGTCGTTCTTCTC 7980
     CTAGCGCTTGTCGTGCTAGCTCTAATGGCGCTTCTACTGCGGTATCTCCAGCAAGAAGAG
```

FIG. 24H

```
7981  GAGCGCGTGGTCGTGATCCTCGTCGTGGTGCGATTTATGCCAGCCCCAGCCCAGGGGGCC  8040
      CTCGCGCACCAGCACTAGGAGCAGCACCACGCTAAATACGGTCGGGGTCGGGTCCCCCGG

8041  ATGCATGATCCACCGATGGACGGAGTAGGCCGTCAGCTCCATCGCGGCGACGGTCAGGAT  8100
      TACGTACTAGGTGGCTACCTGCCTCATCCGGCAGTCGAGGTAGCGCCGCTGCCAGTCCTA

8101  GACGGTCAGGATTGCGGCCCAAGTGCTCATGCCGGCCCCTTGCTTGATATGACAGGGAAC  8160
      CTGCCAGTCCTAACGCCGGGTTCACGAGTACGGCCGGGGAACGAACTATACTGTCCCTTG

8161  AGGCTACGCTGCCGCGCGGTGCATGACCAGCCCATCGGGGTGCGACCAAAGGGCATCGCG  8220
      TCCGATGCGACGGCGCGCCACGTACTGGTCGGGTAGCCCCACGCTGGTTTCCCGTAGCGC

8221  TGACATCTGCGTTCAGGGCTCATAGGCGGATCATCCGTGACATTCGCCGCCGAACGCGGC  8280
      ACTGTAGACGCAAGTCCCGAGTATCCGCCTAGTAGGCACTGTAAGCGGCGGCTTGCGCCG

8281  AGGCGCATCACGCGTTCCGTCGCTGGAAATATTAATGTTTTCCCGAAGATGGTCGGGGCG  8340
      TCCGCGTAGTGCGCAAGGCAGCGACCTTTATAATTACAAAAGGGCTTCTACCAGCCCCGC

8341  AGAGGATTCGAACCTCCGACCTACGGTACCCAAAACCGTCGCGCTACCAGGCTGCGCTAC  8400
      TCTCCTAAGCTTGGAGGCTGGATGCCATGGGTTTTGGCAGCGCGATGGTCCGACGCGATG

8401  GCCCCGACTGCGGAAGGCTTTAGCCGATTGTTCCGGCAAGGGAAAGACCTAGTCGCAGGC  8460
      CGGGGCTGACGCCTTCCGAAATCGGCTAACAAGGCCGTTCCCTTTCTGGATCAGCGTCCG

8461  CAGGACCGCATTGTCGCCCATGCCCGGATGCGCCATCGGCTGACCGGGCTTCAGGCCAAG  8520
      GTCCTGGCGTAACAGCGGGTACGGGCCTACGCGGTAGCCGACTGGCCCGAAGTCCGGTTC

8521  GCGATCCGCCTCTCCGCCCGCGATTTCGAGGACGAACAGCCGGTCGGGGTCCGGATCGCC  8580
      CGCTAGGCGGAGAGGCGGGCGCTAAAGCTCCTGCTTGTCGGCCAGCCCCAGGCCTAGCGG

8581  GACCGCCGCGCCCGGAATGGGCGTCTCGTCCAGCGGGCGCGCATTGCGGTGGATGTGGCG  8640
      CTGGCGGCGCGGGCCTTACCCGCAGAGCAGGTCGCCCGCGCGTAACGCCACCTACACCGC

8641  GATGACGCCGGTTTCATCCGCAAAGACCATGTCCAGCGGGATCAGTGTGTTGCGCATCCA  8700
      CTACTGCGGCCAAAGTAGGCGTTTCTGGTACAGGTCGCCCTAGTCACACAACGCGTAGGT

8701  GAAGGACACCGGCTGGGGCGATTCGTAGATGAACAGCATTCCGGTGCCCGCAGGCAGCTC  8760
      CTTCCTGTGGCCGACCCCGCTAAGCATCTACTTGTCGTAAGGCCACGGGCGTCCGTCGAG

8761  CTTGCGGAACATCAGGCCCTGCGCGCGCTCTTCGGGGCTGTCCGCGACCTCGACCCGAAA  8820
      GAACGCCTTGTAGTCCGGGACGCGCGCGAGAAGCCCCGACAGGCGCTGGAGCTGGGCTTT

8821  CCCGAGCGTTTCCGCACCGGTATCGACGACAAGACTGCCGGGCGCGCATTCCACCGCCGC  8880
      GGGCTCGCAAAGGCGTGGCCATAGCTGCTGTTCTGACGGCCCGCGCGTAAGGTGGCGGCG

8881  CGCGGCGGCGGGCATCAGGACCGCAAGAAGCGCTGCGGCCTTACTCGGCCACATGGGCAA  8940
      GCGCCGCCGCCCGTAGTCCTGGCGTTCTTCGCGACGCCGGAATGAGCCGGTGTACCCGTT

8941  GATAGGACTGCTCGGCGCCGAGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCG  9000
      CTATCCTGACGAGCCGCGGCTCTAGGGGGCCCGACGTCCTTAAGCTATAGTTCGAATAGC
```

FIG. 24I

```
9001 ATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTA 9060
     TATGGCAGCTGGAGCTCCCCCCCGGGCCATGGGTCGAAAACAAGGGAAATCACTCCCAAT

9061 ATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTC 9120
     TAACGCGCGAACCGCATTAGTACCAGTATCGACAAAGGACACACTTTAACAATAGGCGAG

9121 ACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGA 9180
     TGTTAAGGTGTGTTGTATGCTCGGCCTTCGTATTTCACATTTCGGACCCCACGGATTACT

9181 GTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTG 9240
     CACTCGATTGAGTGTAATTAACGCAACGCGAGTGACGGGCGAAAGGTCAGCCCTTTGGAC

9241 TCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG 9300
     AGCACGGTCGACGTAATTACTTAGCCGGTTGCGCGCCCCTCTCCGCCAAACGCATAACCC

9301 CGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCG 9360
     GCGAGAAGGCGAAGGAGCGAGTGACTGAGCGACGCGAGCCAGCAAGCCGACGCCGCTCGC

9361 GTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGA 9420
     CATAGTCGAGTGAGTTTCCGCCATTATGCCAATAGGTGTCTTAGTCCCCTATTGCGTCCT

9421 AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTG 9480
     TTCTTGTACACTCGTTTTCCGGTCGTTTTCCGGTCCTTGGCATTTTTCCGGCGCAACGAC

9481 GCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG 9540
     CGCAAAAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTC

9541 AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTC 9600
     TCCACCGCTTTGGGCTGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAG

9601 GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG 9660
     CACGCGAGAGGACAAGGCTGGGACGGCGAATGGCCTATGGACAGGCGGAAAGAGGGAAGC

9661 GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT 9720
     CCTTCGCACCGCGAAAGAGTATCGAGTGCGACATCCATAGAGTCAAGCCACATCCAGCAA

9721 CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC 9780
     GCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCGGAATAGG

9781 GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC 9840
     CCATTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGG

9841 ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG 9900
     TGACCATTGTCCTAATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAACTTCACC

9901 TGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA 9960
     ACCGGATTGATGCCGATGTGATCTTCCTGTCATAAACCATAGACGCGAGACGACTTCGGT
```

FIG. 24J

```
 9961 GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC 10020
      CAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTTTGTTTGGTGGCGACCATCG

10021 GGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGAT 10080
      CCACCAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTA

10081 CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATT 10140
      GGAAACTAGAAAAGATGCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAA

10141 TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGT 10200
      AACCAGTACTCTAATAGTTTTTCCTAGAAGTGGATCTAGGAAAATTTAATTTTTACTTCA

10201 TTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC 10260
      AAATTTAGTTAGATTTCATATATACTCATTTGAACCAGACTGTCAATGGTTACGAATTAG

10261 AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCC 10320
      TCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCAAGTAGGTATCAACGGACTGAGGGG

10321 GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATA 10380
      CAGCACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCGGGGTCACGACGTTACTAT

10381 CCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGG 10440
      GGCGCTCTGGGTGCGAGTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCC

10441 GCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGC 10500
      CGGCTCGCGTCTTCACCAGGACGTTGAAATAGGCGGAGGTAGGTCAGATAATTAACAACG

10501 CGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT 10560
      GCCCTTCGATCTCATTCATCAAGCGGTCAATTATCAAACGCGTTGCAACAACGGTAACGA

10561 ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAA 10620
      TGTCCGTAGCACCACAGTGCGAGCAGCAAACCATACCGAAGTAAGTCGAGGCCAAGGGTT

10621 CGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGT 10680
      GCTAGTTCCGCTCAATGTACTAGGGGGTACAACACGTTTTTTCGCCAATCGAGGAAGCCA

10681 CCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCA 10740
      GGAGGCTAGCAACAGTCTTCATTCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGT

10741 CTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTAC 10800
      GACGTATTAAGAGAATGACAGTACGGTAGGCATTCTACGAAAAGACACTGACCACTCATG

10801 TCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCA 10860
      AGTTGGTTCAGTAAGACTCTTATCACATACGCCGCTGGCTCAACGAGAACGGGCCGCAGT

10861 ATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGT 10920
      TATGCCCTATTATGGCGCGGTGTATCGTCTTGAAATTTTCACGAGTAGTAACCTTTTGCA

10921 TCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCC 10980
      AGAAGCCCCGCTTTTGAGAGTTCCTAGAATGGCGACAACTCTAGGTCAAGCTACATTGGG
```

FIG. 24K

```
      ACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCA
10981 ---------+---------+---------+---------+---------+---------+ 11040
      TGAGCACGTGGGTTGACTAGAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGT

      AAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATA
11041 ---------+---------+---------+---------+---------+---------+ 11100
      TTTTGTCCTTCCGTTTTACGGCGTTTTTTCCCTTATTCCCGCTGTGCCTTTACAACTTAT

      CTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGC
11101 ---------+---------+---------+---------+---------+---------+ 11160
      GAGTATGAGAAGGAAAAAGTTATAATAACTTCGTAAATAGTCCCAATAACAGAGTACTCG

      GGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC
11161 ---------+---------+---------+---------+---------+---------+ 11220
      CCTATGTATAAACTTACATAAATCTTTTTATTTGTTTATCCCCAAGGCGCGTGTAAAGGG

      CGAAAAGTGCCAC
11221 ---------+--- 11233
      GCTTTTCACGGTG
```

FIG. 24L

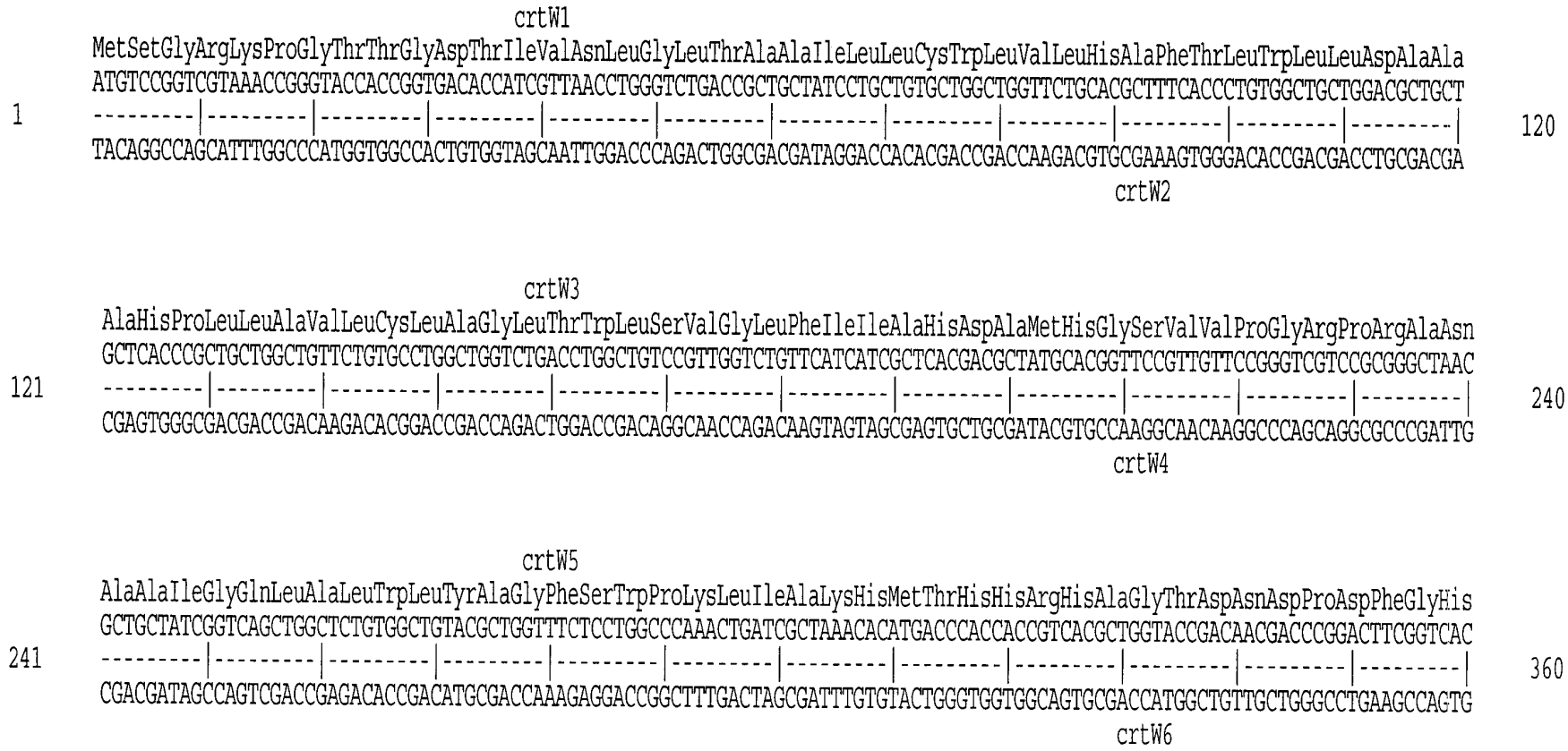

crtW1
MetSetGlyArgLysProGlyThrThrGlyAspThrIleValAsnLeuGlyLeuThrAlaAlaIleLeuLeuCysTrpLeuValLeuHisAlaPheThrLeuTrpLeuLeuAspAlaAla
ATGTCCGGTCGTAAACCGGGTACCACCGGTGACACCATCGTTAACCTGGGTCTGACCGCTGCTATCCTGCTGTGCTGGCTGGTTCTGCACGCTTTCACCCTGTGGCTGCTGGACGCTGCT
1
120
TACAGGCCAGCATTTGGCCCATGGTGGCCACTGTGGTAGCAATTGGACCCAGACTGGCGACGATAGGACCACACGACCGACCAAGACGTGCGAAAGTGGGACACCGACGACCTGCGACGA
crtW2
crtW3
AlaHisProLeuLeuAlaValLeuCysLeuAlaGlyLeuThrTrpLeuSerValGlyLeuPheIleIleAlaHisAspAlaMetHisGlySerValValProGlyArgProArgAlaAsn
GCTCACCCGCTGCTGGCTGTTCTGTGCCTGGCTGGTCTGACCTGGCTGTCCGTTGGTCTGTTCATCATCGCTCACGACGCTATGCACGGTTCCGTTGTTCCGGGTCGTCCGCGGGCTAAC
121
240
CGAGTGGGCGACGACCGACAAGACACGGACCGACCAGACTGGACCGACAGGCAACCAGACAAGTAGTAGCGAGTGCTGCGATACGTGCCAAGGCAACAAGGCCCAGCAGGCGCCCGATTG
crtW4
crtW5
AlaAlaIleGlyGlnLeuAlaLeuTrpLeuTyrAlaGlyPheSerTrpProLysLeuIleAlaLysHisMetThrHisHisArgHisAlaGlyThrAspAsnAspProAspPheGlyHis
GCTGCTATCGGTCAGCTGGCTCTGTGGCTGTACGCTGGTTTCTCCTGGCCCAAACTGATCGCTAAACACATGACCCACCACCGTCACGCTGGTACCGACAACGACCCGGACTTCGGTCAC
241
360
CGACGATAGCCAGTCGACCGAGACACCGACATGCGACCAAAGAGGACCGGCTTTGACTAGCGATTTGTGTACTGGGTGGTGGCAGTGCGACCATGGCTGTTGCTGGGCCTGAAGCCAGTG
crtW6

FIG. 25A

crtW7
GlyGlyProValArgTrpTyrGlySerPheValSerThrTyrPheGlyTrpArgGluGlyLeuLeuLeuProValIleValThrThrTyrAlaLeuIleLeuGlyAspArgTrpMetTyr
GGTGGTCCGGTTCGTTGGTACGGTTCCTTCGTTTCCACCTACTTCGGTTGGCGTGAAGGTCTGCTGCTGCCGGTTATCGTACCACCTACGCTCTGATCCTGGGTGACCGTTGGATGTAC
361
480
CCACCAGGCCAAGCAACCATGCCAAGGAAGCAAAGGTGGATGAAGCCAACCGCACTTCCAGACGACGACGGCCAATAGCAATGGTGGATGCGAGACTAGGACCCACTGGCAACCTACATG
crtW8
crtW9
ValIlePheTrpProValProAlaValLeuAlaSerIleGlnIlePheValPheGlyThrTrpLeuProHisArgProGlyHisAspAspPheProAspArgHisAsnAlaArgSerThr
GTTATCTTCTGGCCGGTTCCGGCTGTTCTGGCTTCCATCCAGATCTTCGTTTTCGGTACCTGGCTGCCGCACCGTCCGGGTCACGACGACTTCCCGGACCGTCACAACGCTCGTTCCACC
481
600
CAATAGAAGACCGGCCAAGGCCGACAAGACCGAAGGTAGGTCTAGAAGCAAAAGCCATGGACCGACGGCGTGGCAGGCCCAGTGCTGCTGAAGGGCCTGGCAGYGTTGCGAGCAAGGTGG
crtW10
crtW11
GlyIleGlyAspProLeuSerLeuLeuThrCysPheHisPheGlyGlyTyrHisHisGluHisHisLeuHisProHisValProTrpTrpArgLeuProArgThrArgLysThrGlyGly
GGTATCGGTGACCCGCTGTCCCTGCTGACCTGCTTCCACTTCGGTGGTTACCACCACGAACACCACCTGCACCCGCACGTTCCGTGGTGGCGTCTGCCGCGTACCCGTAAAACCGGTGGT
601
720
CCATAGCCACTGGGCGACAGGGACGACTGGACGAAGGTGAAGCCACCAATGGTGGTGCTTGTGGTGGACGTGGGCGTGCAAGGCACCACCGCAGACGGCGCATGGGCATTTTGGCCACCA
crtW12
ArgAla
CGTGCT
721
726
GCACGA

FIG. 25B

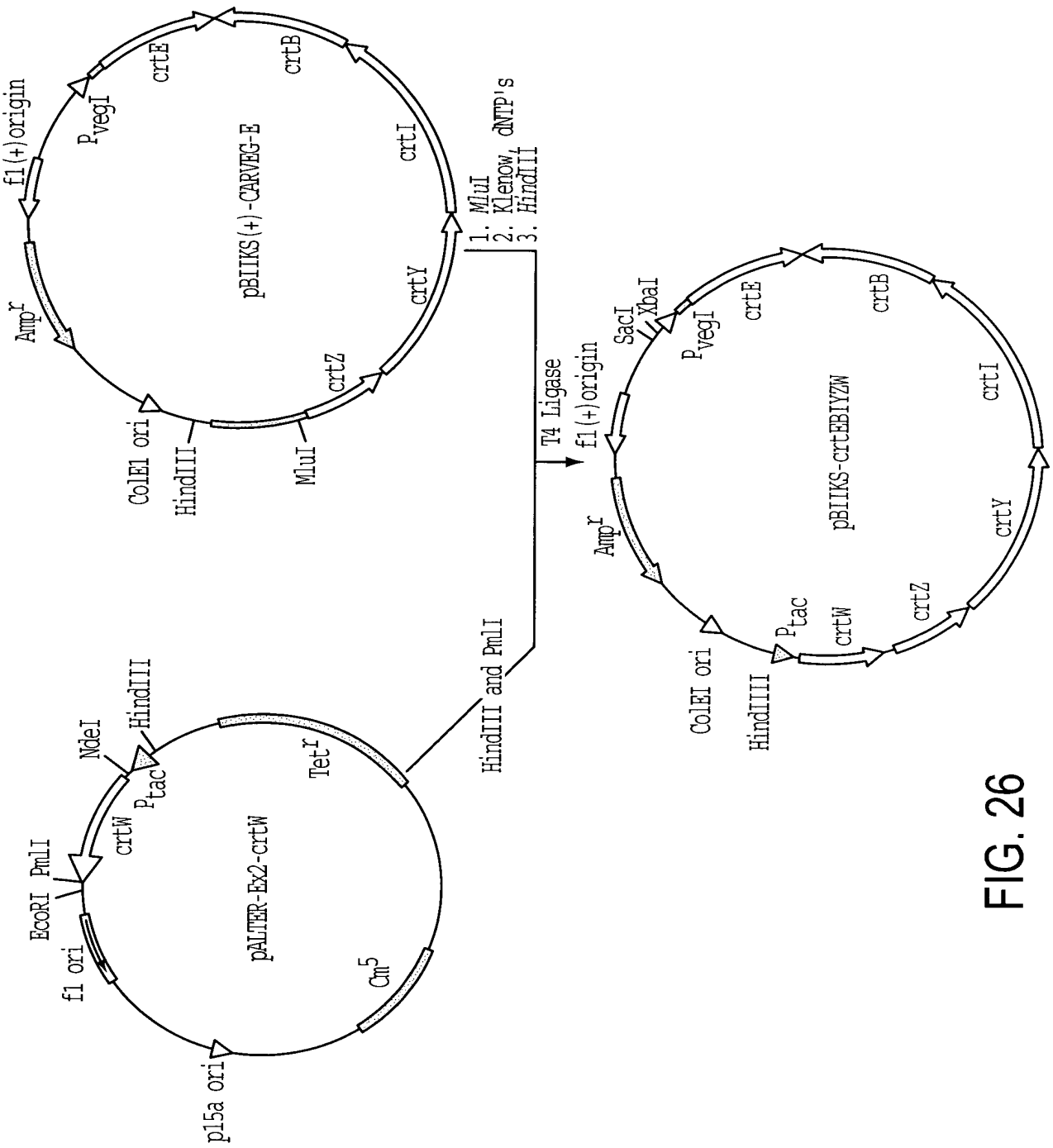

pBIIKS(+)-CARVEG-E
f1(+)origin
P vegI
crtE
crtB
crtI
crtY
crtZ
Amp r
ColE1 ori
HindIII
MluI
1. MluI
2. Klenow, dNTP's
3. HindIII
T4 Ligase
HindIII and PmlI
pALTER-Ex2-crtW
NdeI
HindIII
P tac
crtW
Tet r
EcoRI
PmlI
f1 ori
Cm 5
p15a ori
pBIIKS-crtEBIYZW
SacI
XbaI
f1(+)origin
P vegI
crtE
crtB
crtI
crtY
crtZ
Amp r
ColEI ori
HindIIII
P tac
crtW

FIG. 26

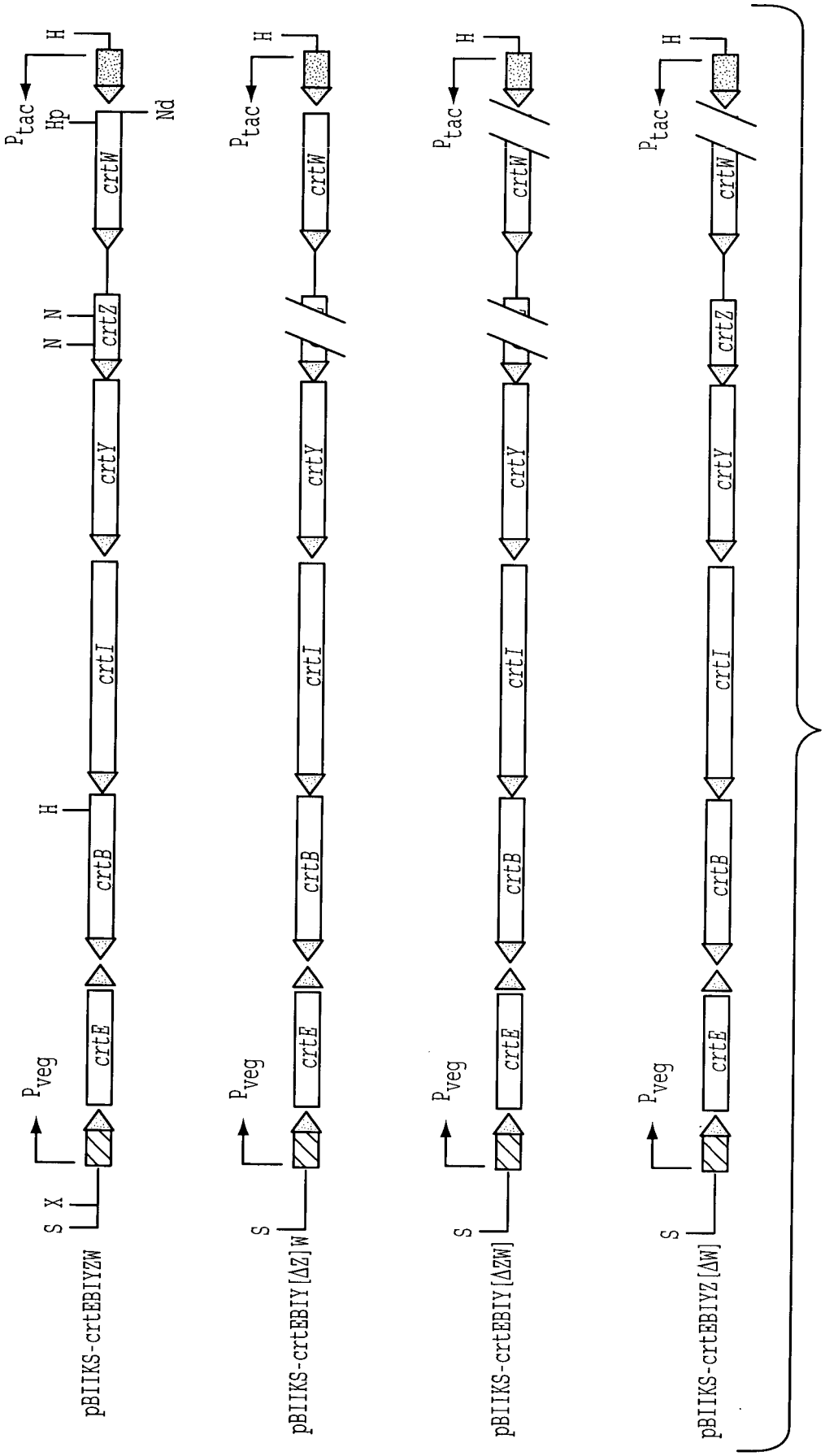

pBIIKS-crtEBIYZW
pBIIKS-crtEBIY[ΔZ]W
pBIIKS-crtEBIY[ΔZW]
pBIIKS-crtEBBIYZ[ΔW]
S
X
P veg
crtE
crtB
H
crtI
crtY
N
N
crtZ
Hp
Nd
P tac
crtW

FIG. 27

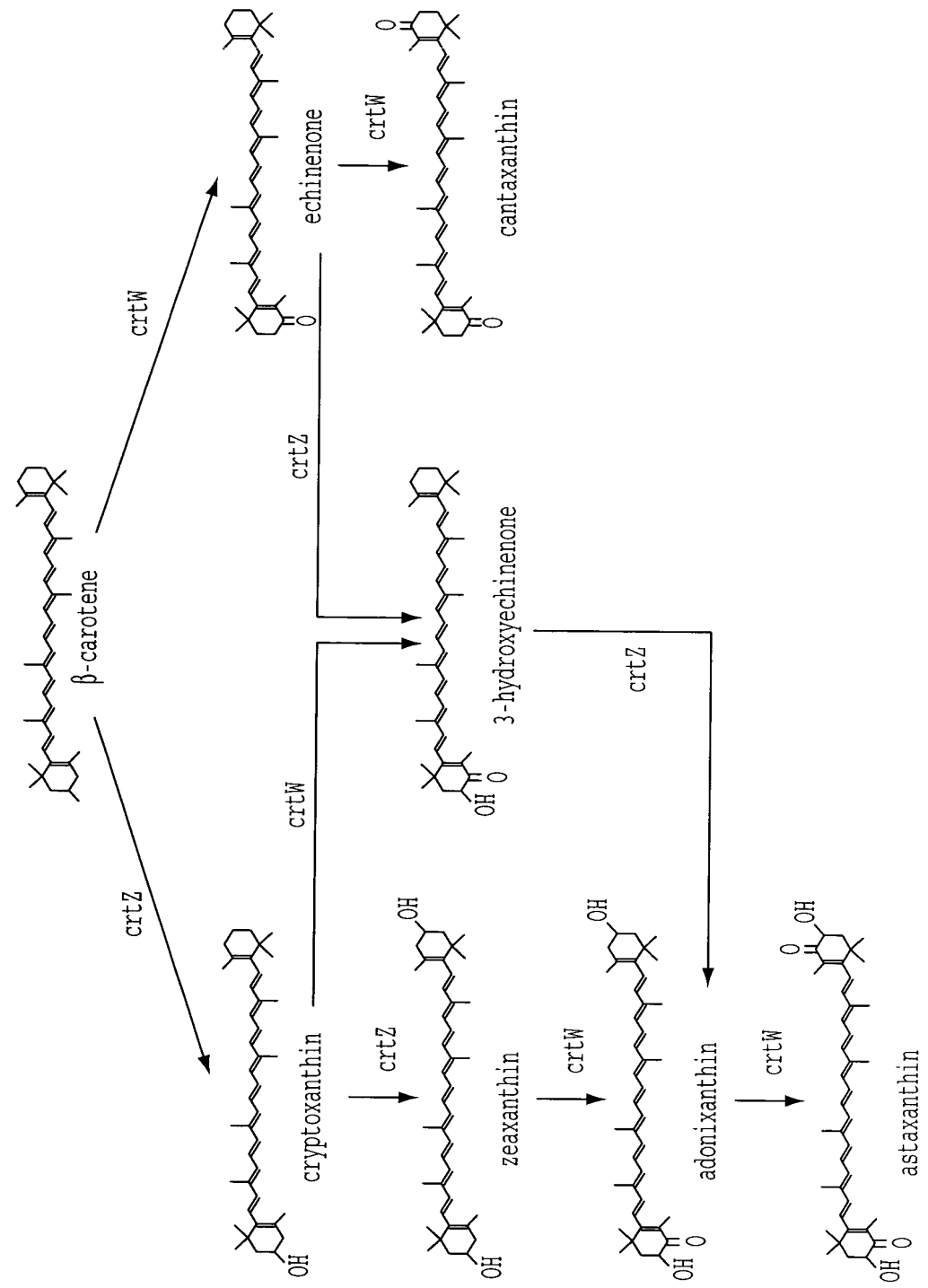

β-carotene
crtW
crtZ
echinenone
crtW
cantaxanthin
crtZ
3-hydroxyechinenone
crtW
crtZ
cryptoxanthin
crtZ
zeaxanthin
crtW
adonixanthin
crtW
astaxanthin
OH
O

FIG. 28

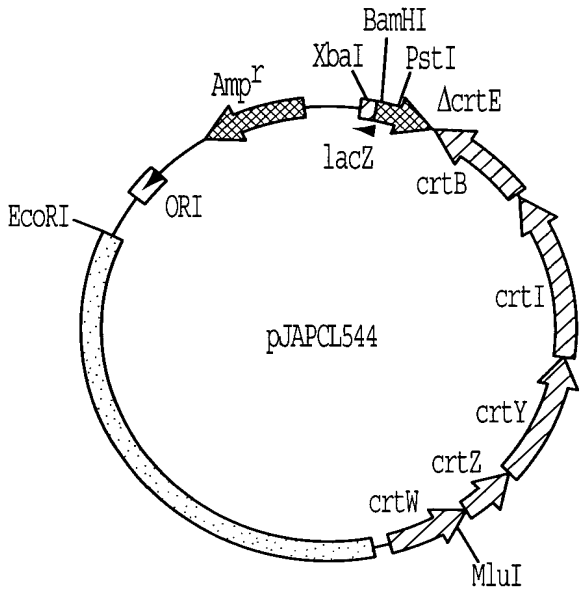

BamHI
XbaI
PstI
Amp^r
ΔcrtE
lacZ
crtB
EcoRI
ORI
crtI
pJAPCL544
crtY
crtZ
crtW
MluI

FIG. 29

```
      ACTGTAGTCTGCGCGGATCGCCGGTCCGGGGGACAAGATATGAGCGCACATGCCCTGCCC
    1 ---------+---------+---------+---------+---------+---------+   60
      TGACATCAGACGCGCCTAGCGGCCAGGCCCCCTGTTCTATACTCGCGTGTACGGGACGGG

      AAGGCAGATCTGACCGCCACCAGTTTGATCGTCTCGGGCGGCATCATCGCCGCGTGGCTG
   61 ---------+---------+---------+---------+---------+---------+  120
      TTCCGTCTAGACTGGCGGTGGTCAAACTAGCAGAGCCCGCCGTAGTAGCGGCGCACCGAC

      GCCCTGCATGTGCATGCGCTGTGGTTTCTGGACGCGGCGGCGCATCCCATCCTGGCGGTC
  121 ---------+---------+---------+---------+---------+---------+  180
      CGGGACGTACACGTACGCGACACCAAAGACCTGCGCCGCCGCGTAGGGTAGGACCGCCAG

      GCGAATTTCCTGGGGCTGACCTGGCTGTCGGTCGGTCTGTTCATCATCGCGCATGACGCG
  181 ---------+---------+---------+---------+---------+---------+  240
      CGCTTAAAGGACCCCGACTGGACCGACAGCCAGCCAGACAAGTAGTAGCGCGTACTGCGC

      ATGCATGGGTCGGTCGTGCCGGGGCGCCCGCGCGCCAATGCGGCGATGGGCCAGCTTGTC
  241 ---------+---------+---------+---------+---------+---------+  300
      TACGTACCCAGCCAGCACGGCCCCGCGGGCGCGCGGTTACGCCGCTACCCGGTCGAACAG

      CTGTGGCTGTATGCCGGATTTTCCTGGCGCAAGATGATCGTCAAGCACATGGCCCATCAT
  301 ---------+---------+---------+---------+---------+---------+  360
      GACACCGACATACGGCCTAAAAGGACCGCGTTCTACTAGCAGTTCGTGTACCGGGTAGTA

      CGCCATGCCGGAACCGACGACGACCCAGATTTCGACCATGGCGGCCCGGTCCGCTGGTAC
  361 ---------+---------+---------+---------+---------+---------+  420
      GCGGTACGGCCTTGGCTGCTGCTGGGTCTAAAGCTGGTACCGCCGGGCCAGGCGACCATG

      GCCCGCTTCATCGGCACCTATTTCGGCTGGCGCGAGGGGCTGCTGCTGCCCGTCATCGTG
  421 ---------+---------+---------+---------+---------+---------+  480
      CGGGCGAAGTAGCCGTGGATAAAGCCGACCGCGCTCCCCGACGACGACGGGCAGTAGCAC

      ACGGTCTATGCGCTGATGTTGGGGGATCGCTGGATGTACGTGGTCTTCTGGCCGTTGCCG
  481 ---------+---------+---------+---------+---------+---------+  540
      TGCCAGATACGCGACTACAACCCCCTAGCGACCTACATGCACCAGAAGACCGGCAACGGC

      TCGATCCTGGCGTCGATCCAGCTGTTCGTGTTCGGCATCTGGCTGCCGCACCGCCCCGGC
  541 ---------+---------+---------+---------+---------+---------+  600
      AGCTAGGACCGCAGCTAGGTCGACAAGCACAAGCCGTAGACCGACGGCGTGGCGGGGCCG

      CACGACGCGTTCCCGGACCGCCACAATGCGCGGTCGTCGCGGATCAGCGACCCCGTGTCG
  601 ---------+---------+---------+---------+---------+---------+  660
      GTGCTGCGCAAGGGCCTGGCGGTGTTACGCGCCAGCAGCGCCTAGTCGCTGGGGCACAGC
```

FIG. 30A

```
      CTGCTGACCTGCTTTCACTTTGGCGGTTATCATCACGAACACCACCTGCACCCGACGGTG
661   ---------+---------+---------+---------+---------+---------+   720
      GACGACTGGACGAAAGTGAAACCGCCAATAGTAGTGCTTGTGGTGGACGTGGGCTGCCAC

      CCTTGGTGGCGCCTGCCCAGCACCCGCACCAAGGGGGACACCGCATGACCAATTTCCTGA
721   ---------+---------+---------+---------+---------+---------+   780
      GGAACCACCGCGGACGGGTCGTGGGCGTGGTTCCCCCTGTGGCGTACTGGTTAAAGGACT

      TCGTCGTCGCCACCGTGCTGGTGATGGAGCTGACGGCCTATTCCGTCCACCGCTGGATCA
781   ---------+---------+---------+---------+---------+---------+   840
      AGCAGCAGCGGTGGCACGACCACTACCTCGACTGCCGGATAAGGCAGGTGGCGACCTAGT

      TGCACGGCCCCTTGGGCTGGGGCTGGCACAAGTCCCACCACGAGGAACACGACCACGCGC
841   ---------+---------+---------+---------+---------+---------+   900
      ACGTGCCGGGGAACCCGACCCCGACCGTGTTCAGGGTGGTGCTCCTTGTGCTGGTGCGCG

      TGGAAAAGAACGACCTGTACGGCCTGGTCTTTGCGGTGATCGCCACGGTGCTGTTCACGG
901   ---------+---------+---------+---------+---------+---------+   960
      ACCTTTTCTTGCTGGACATGCCGGACCAGAAACGCCACTAGCGGTGCCACGACAAGTGCC

      TGGGCTGGATCTGGGCACCGGTCCTGTGGTGGATCGCCTTGGGCATGACCGTCTACGGGC
961   ---------+---------+---------+---------+---------+---------+   1020
      ACCCGACCTAGACCCGTGGCCAGGACACCACCTAGCGGAACCCGTACTGGCAGATGCCCG

      TGATCTATTTCGTCCTGCATGACGGGCTGGTGCATCAGCGCTGGCCGTTCCGCTATATCC
1021  ---------+---------+---------+---------+---------+---------+   1080
      ACTAGATAAAGCAGGACGTACTGCCCGACCACGTAGTCGCGACCGGCAAGGCGATATAGG

      CTCGCAAGGGCTATGCCAGACGCCTGTATCAGGCCCACCGCCTGCACCACGCGGTCGAGG
1081  ---------+---------+---------+---------+---------+---------+   1140
      GAGCGTTCCCGATACGGTCTGCGGACATAGTCCGGGTGGCGGACGTGGTGCGCCAGCTCC

      GGCGCGACCATTGCGTCAGCTTCGGCTTCATCTATGCGCCGCCGGTCGACAAGCTGAAGC
1141  ---------+---------+---------+---------+---------+---------+   1200
      CCGCGCTGGTAACGCAGTCGAAGCCGAAGTAGATACGCGGCGGCCAGCTGTTCGACTTCG

      AGGACCTGAAGACGTCGGGCGTGCTGCGGGCCGAGGCGCAGGAGCGCACGTGACCCATGA
1201  ---------+---------+---------+---------+---------+---------+   1260
      TCCTGGACTTCTGCAGCCCGCACGACGCCCGGCTCCGCGTCCTCGCGTGCACTGGGTACT

      C
1261  -  1261
      G
```

FIG. 30B

```
    ATGAGCGCACATGCCCTGCCCAAGGCAGATCTGACCGCCACCAGTTTGATCGTCTCGGGC
  1 ---------+---------+---------+---------+---------+---------+   60
    TACTCGCGTGTACGGGACGGGTTCCGTCTAGACTGGCGGTGGTCAAACTAGCAGAGCCCG

    GGCATCATCGCCGCGTGGCTGGCCCTGCATGTGCATGCGCTGTGGTTTCTGGACGCGGCG
 61 ---------+---------+---------+---------+---------+---------+  120
    CCGTAGTAGCGGCGCACCGACCGGGACGTACACGTACGCGACACCAAAGACCTGCGCCGC

    GCGCATCCCATCCTGGCGGTCGCGAATTTCCTGGGGCTGACCTGGCTGTCGGTCGGTCTG
121 ---------+---------+---------+---------+---------+---------+  180
    CGCGTAGGGTAGGACCGCCAGCGCTTAAAGGACCCCGACTGGACCGACAGCCAGCCAGAC

    TTCATCATCGCGCATGACGCGATGCATGGGTCGGTCGTGCCGGGGCGCCCGCGCGCCAAT
181 ---------+---------+---------+---------+---------+---------+  240
    AAGTAGTAGCGCGTACTGCGCTACGTACCCAGCCAGCACGGCCCCGCGGGCGCGCGGTTA

    GCGGCGATGGGCCAGCTTGTCCTGTGGCTGTATGCCGGATTTTCCTGGCGCAAGATGATC
241 ---------+---------+---------+---------+---------+---------+  300
    CGCCGCTACCCGGTCGAACAGGACACCGACATACGGCCTAAAAGGACCGCGTTCTACTAG

    GTCAAGCACATGGCCCATCATCGCCATGCCGGAACCGACGACGACCCAGATTTCGACCAT
301 ---------+---------+---------+---------+---------+---------+  360
    CAGTTCGTGTACCGGGTAGTAGCGGTACGGCCTTGGCTGCTGCTGGGTCTAAAGCTGGTA

    GGCGGCCCGGTCCGCTGGTACGCCCGCTTCATCGGCACCTATTTCGGCTGGCGCGAGGGG
361 ---------+---------+---------+---------+---------+---------+  420
    CCGCCGGGCCAGGCGACCATGCGGGCGAAGTAGCCGTGGATAAAGCCGACCGCGCTCCCC

    CTGCTGCTGCCCGTCATCGTGACGGTCTATGCGCTGATGTTGGGGGATCGCTGGATGTAC
421 ---------+---------+---------+---------+---------+---------+  480
    GACGACGACGGGCAGTAGCACTGCCAGATACGCGACTACAACCCCCTAGCGACCTACATG

    GTGGTCTTCTGGCCGTTGCCGTCGATCCTGGCGTCGATCCAGCTGTTCGTGTTCGGCATC
481 ---------+---------+---------+---------+---------+---------+  540
    CACCAGAAGACCGGCAACGGCAGCTAGGACCGCAGCTAGGTCGACAAGCACAAGCCGTAG

    TGGCTGCCGCACCGCCCCGGCCACGACGCGTTCCCGGACCGCCACAATGCGCGGTCGTCG
541 ---------+---------+---------+---------+---------+---------+  600
    ACCGACGGCGTGGCGGGGCCGGTGCTGCGCAAGGGCCTGGCGGTGTTACGCGCCAGCAGC

    CGGATCAGCGACCCCGTGTCGCTGCTGACCTGCTTTCACTTTGGCGGTTATCATCACGAA
601 ---------+---------+---------+---------+---------+---------+  660
    GCCTAGTCGCTGGGGCACAGCGACGACTGGACGAAAGTGAAACCGCCAATAGTAGTGCTT

    CACCACCTGCACCCGACGGTGCCTTGGTGGCGCCTGCCCAGCACCCGCACCAAGGGGGAC
661 ---------+---------+---------+---------+---------+---------+  720
    GTGGTGGACGTGGGCTGCCACGGAACCACCGCGGACGGGTCGTGGGCGTGGTTCCCCCTG

    ACCGCATGA
721 ---------  729
    TGGCGTACT
```

FIG. 31

```
  1  MSAHALPKAD LTATSLIVSG GIIAAWLALH VHALWFLDAA AHPILAVANF

 51  LGLTWLSVGL FIIAHDAMHG SVVPGRPRAN AAMGQLVLWL YAGFSWRKMI

101  VKHMAHHRHA GTDDDPDFDH GGPVRWYARF IGTYFGWREG LLLPVIVTVY

151  ALMLGDRWMY VVFWPLPSIL ASIQLFVFGI WLPHRPGHDA FPDRHNARSS

201  RISDPVSLLT CFHFGGYHHE HHLHPTVPWW RLPSTRTKGD TA*
```

FIG. 32

```
    ATGACCAATTTCCTGATCGTCGTCGCCACCGTGCTGGTGATGGAGCTGACGGCCTATTCC
  1 ---------+---------+---------+---------+---------+---------+   60
    TACTGGTTAAAGGACTAGCAGCAGCGGTGGCACGACCACTACCTCGACTGCCGGATAAGG

    GTCCACCGCTGGATCATGCACGGCCCCTTGGGCTGGGGCTGGCACAAGTCCCACCACGAG
 61 ---------+---------+---------+---------+---------+---------+  120
    CAGGTGGCGACCTAGTACGTGCCGGGGAACCCGACCCCGACCGTGTTCAGGGTGGTGCTC

    GAACACGACCACGCGCTGGAAAAGAACGACCTGTACGGCCTGGTCTTTGCGGTGATCGCC
121 ---------+---------+---------+---------+---------+---------+  180
    CTTGTGCTGGTGCGCGACCTTTTCTTGCTGGACATGCCGGACCAGAAACGCCACTAGCGG

    ACGGTGCTGTTCACGGTGGGCTGGATCTGGGCACCGGTCCTGTGGTGGATCGCCTTGGGC
181 ---------+---------+---------+---------+---------+---------+  240
    TGCCACGACAAGTGCCACCCGACCTAGACCCGTGGCCAGGACACCACCTAGCGGAACCCG

    ATGACCGTCTACGGGCTGATCTATTTCGTCCTGCATGACGGGCTGGTGCATCAGCGCTGG
241 ---------+---------+---------+---------+---------+---------+  300
    TACTGGCAGATGCCCGACTAGATAAAGCAGGACGTACTGCCCGACCACGTAGTCGCGACC

    CCGTTCCGCTATATCCCTCGCAAGGGCTATGCCAGACGCCTGTATCAGGCCCACCGCCTG
301 ---------+---------+---------+---------+---------+---------+  360
    GGCAAGGCGATATAGGGAGCGTTCCCGATACGGTCTGCGGACATAGTCCGGGTGGCGGAC

    CACCACGCGGTCGAGGGGCGCGACCATTGCGTCAGCTTCGGCTTCATCTATGCGCCGCCG
361 ---------+---------+---------+---------+---------+---------+  420
    GTGGTGCGCCAGCTCCCCGCGCTGGTAACGCAGTCGAAGCCGAAGTAGATACGCGGCGGC

    GTCGACAAGCTGAAGCAGGACCTGAAGACGTCGGGCGTGCTGCGGGCCGAGGCGCAGGAG
421 ---------+---------+---------+---------+---------+---------+  480
    CAGCTGTTCGACTTCGTCCTGGACTTCTGCAGCCCGCACGACGCCCGGCTCCGCGTCCTC

    CGCACG
481 ------  486
    GCGTGC
```

FIG. 33

```
  1  MTNFLIVVAT VLVMELTAYS VHRWIMHGPL GWGWHKSHHE EHDHALEKND

 51  LYGLVFAVIA TVLFTVGWIW APVLWWIALG MTVYGLIYFV LHDGLVHQRW

101  PFRYIPRKGY ARRLYQAHRL HHAVEGRDHC VSFGFIYAPP VDKLKQDLKT

151  SGVLRAEAQE RT
```

FIG. 34

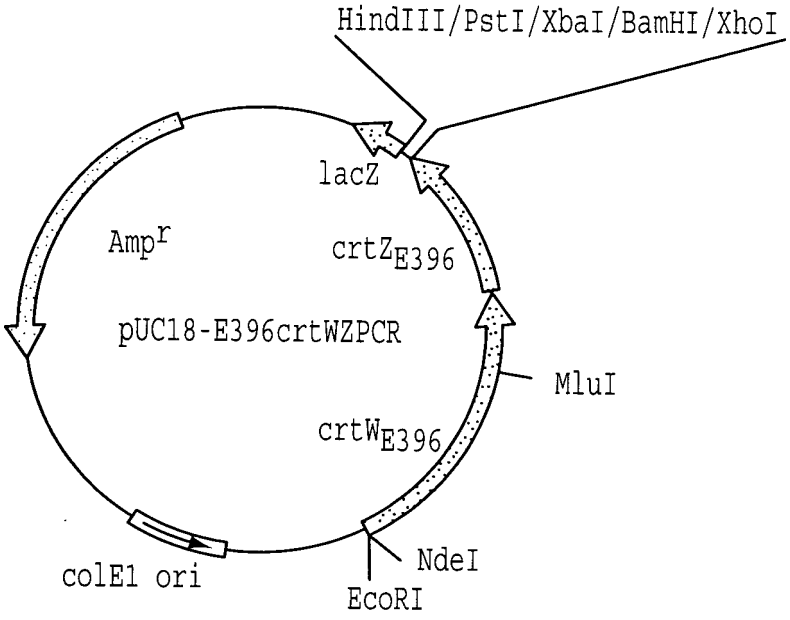

HindIII/PstI/XbaI/BamHI/XhoI
lacZ
Amp^r
crtZ$_{E396}$
pUC18-E396crtWZPCR
MluI
crtW$_{E396}$
colE1 ori
NdeI
EcoRI

FIG. 35

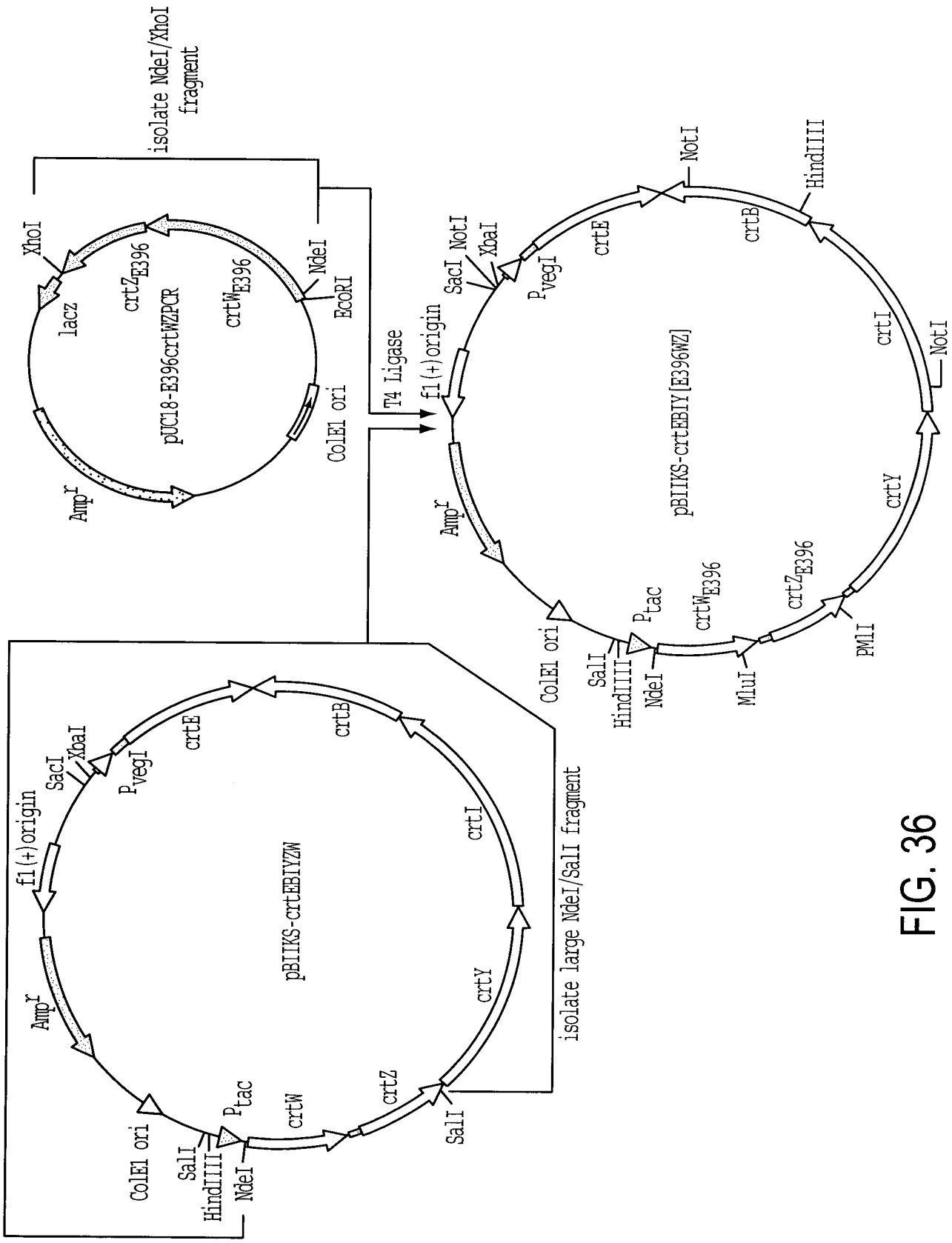

isolate NdeI/XhoI fragment
XhoI
lacz
crtZ E396
crtW E396
NdeI
EcoRI
pUC18-E396crtWZPCR
ColE1 ori
Amp r
T4 Ligase
SacI NotI XbaI
f1(+)origin
P vegI
crtE
NotI
HindIIII
crtB
crtI
crtY
pBIIKS-crtEBIY[E396WZ]
ColE1 ori
SalI
HindIIII
NdeI
P tac
crtW E396
MluI
crtZ E396
PMlI
f1(+)origin
SacI
XbaI
P vegI
crtE
crtB
crtI
pBIIKS-crtEBIYZW
Amp r
ColE1 ori
SalI
HindIIII
NdeI
P tac
crtW
crtZ
crtY
SalI
isolate large NdeI/SalI fragment

FIG. 36

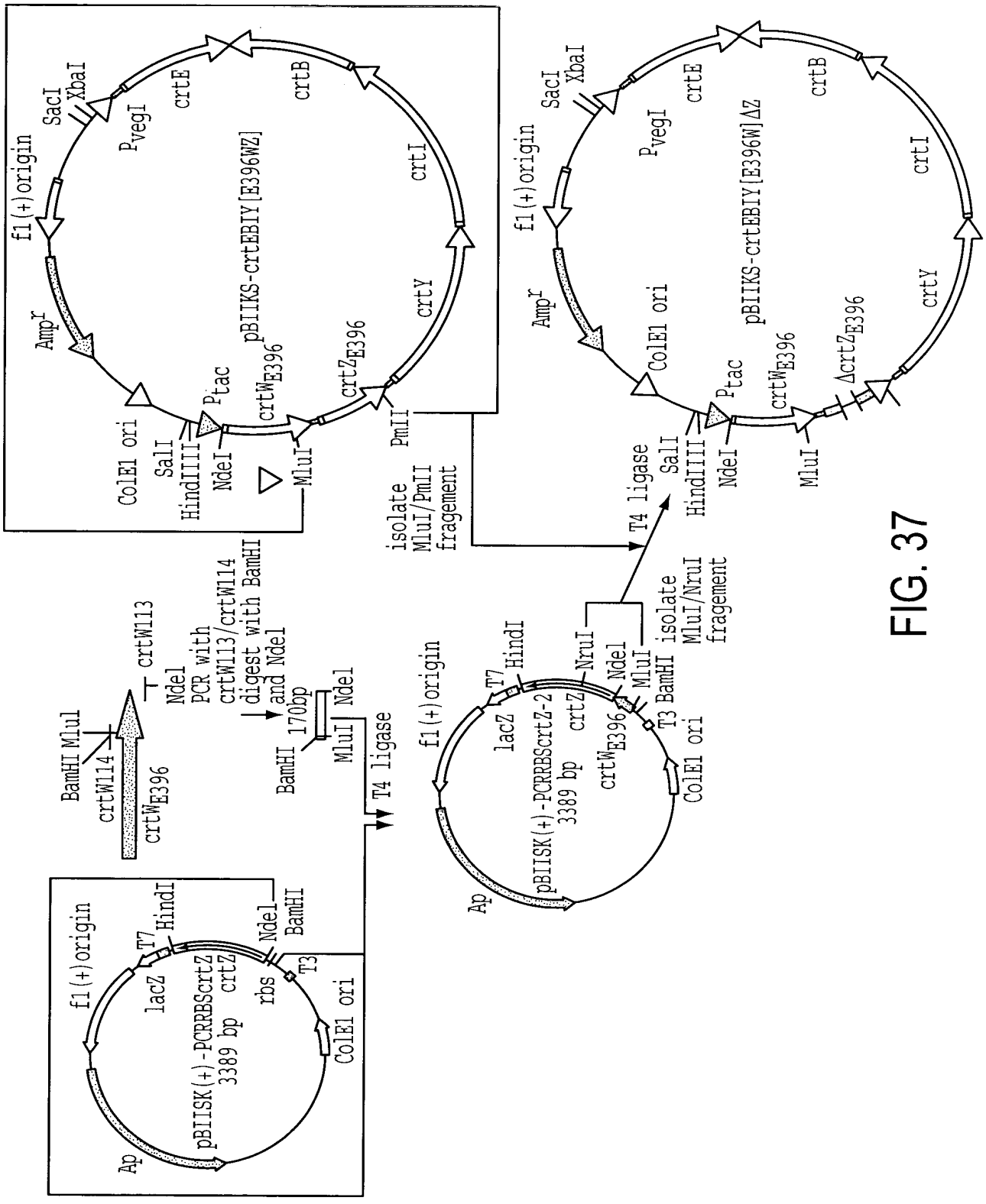

f1(+)origin
SacI
XbaI
P_{vegI}
crtE
crtB
crtI
crtY
Amp^r
ColE1 ori
SalI
HindIIII
P_{tac}
NdeI
$crtW_{E396}$
MluI
$crtZ_{E396}$
PmlI
pBIIKS-crtEBIY[E396WZ]
isolate MluI/PmlI fragement
T4 ligase
pBIIKS-crtEBIY[E396W]ΔZ
$\Delta crtZ_{E396}$
BamHI MluI
crtW114
$crtW_{E396}$
crtW113
NdeI
PCR with crtW113/crtW114
digest with BamHI and NdeI
BamHI 170bp NdeI
MluI
pBIISK(+)-PCRRBScrtZ-2
3389 bp
lacZ
T7
HindI
crtZ
NruI
rbs
T3
BamHI
isolate MluI/NruI fragement
Ap

FIG. 37

```
      CTGCAGGTCTGACACGGCCAGAAGGCCGCGCCGCGGGcCGGGGGCCGCcGCATCGCGACC
1     --------+--------+--------+--------+--------+--------+--------+    60
      GACGTCCAGACTGTGCCGGTCTTCCGGCGCGGCGCCCgGCCCCCGGCGgCGTAGCGCTGG

      GGTATCCTTGCCAAGCGCCGCCTGGTCGCCCACaACGTCCAGCAGGTCGTCATAGGACTG
61    --------+--------+--------+--------+--------+--------+--------+   120
      CCATAGGAACGGTTCGCGGCGGACCAGCGGGTGtTGCAGGTCGTCCAGCAGTATCCTGAC

      GAACACCCGGCCCAGCTGACGGCCAAAGTCGATCATCTGaGTCTGCTCCTCGGCGTCGAA
121   --------+--------+--------+--------+--------+--------+--------+   180
      CTTGTGGGCCGGGTCGACTGCCGGTTTCAGCTAGTAGACtCAGACGAGGAGCCGCAGCTT

      CTCCTTGATCACGGCCAGCATCTCCAGCCCGGCGATGAACAGCACGCCGGTCTTCAGGTC
181   --------+--------+--------+--------+--------+--------+--------+   240
      GAGGAACTAGTGCCGGTCGTAGAGGTCGGGCCGCTACTTGTCGTGCGGCCAGAAGTCCAG

      CTGTTCCTGTTCGACCCCCGCGCCGTTCTTGGCCGCGTGCAGGTCCAGGTCCTGGCCGGC
241   --------+--------+--------+--------+--------+--------+--------+   300
      GACAAGGACAAGCTGGGGGCGCGGCAAGAACCGGCGCACGTCCAGGTCCAGGACCGGCCG

      GCACAGGCCCTGCGGCCCCAGGGACCGCGACAGGATCCgcaccagctgcgcccgcaccgt
301   --------+--------+--------+--------+--------+--------+--------+   360
      CGTGTCCGGGACGCCGGGGTCCCTGGCGCTGTCCTAGGcgtggtcgacgcgggcgtggca

      gcccgacgcgccgcgcgcaccggccagcagggccatcgcctcggtgatcagggcgatgcc
361   --------+--------+--------+--------+--------+--------+--------+   420
      cgggctgcgcggcgcgcgtggccggtcgtcccggtagcggagccactagtcccgctacgg

      gcctagcacggcgcggctttcgccatgcgccacatgggtcgcgggctggccgcggcgcag
421   --------+--------+--------+--------+--------+--------+--------+   480
      cggatcgtgccgcgccgaaagcggtacgcggtgtacccagcgcccgaccggcgccgcgtc

      cccggcatcgtccatgcagggcaggtcgtcgaagatcagcgatgcggcatgcaccatctc
481   --------+--------+--------+--------+--------+--------+--------+   540
      gggccgtagcaggtacgtcccgtccagcagcttctagtcgctacgccgtacgtggtagag

      gaccgcgcaggcggcgtcgacgatcgtgtcgcagaccccgcccgaggcttctgccgcaag
541   --------+--------+--------+--------+--------+--------+--------+   600
      ctggcgcgtccgccgcagctgctagcacagcgtctggggcgggctccgaagacggcgttc

      cagcatcagcatgccgcggaaacgcttgcccgacgacagcgcgccatggctcatggccgg
601   --------+--------+--------+--------+--------+--------+--------+   660
      gtcgtagtcgtacggcgcctttgcgaacgggctgctgtcgcgcggtaccgagtaccggcc

      gccgagcggctgcgacacggcaccgaatccctgggcgatctcctcaagtctggtctgcag
661   --------+--------+--------+--------+--------+--------+--------+   720
      cggctcgccgacgctgtgccgtggcttagggacccgctagaggagttcagaccagacgtc
```

FIG. 38A

```
                aagggtggcgtggatcgggttgacgtctcgtctcatcagtgccttcgcgcttgggttctg
721             --------+--------+--------+--------+--------+--------+        780
                ttcccaccgcacctaggggaactgcagagcagagtagtcacggaagcgcgaacccaagac

                accaggcgggaaggtcaggccggggcggcaccccgtgacccgtcatccaccgtcaacagt
781             --------+--------+--------+--------+--------+--------+        840
                tggtccgcccttccagtccggccccgccgtggggcactgggcagtaggtggcagttgtca

                ccccatgttggaaggcttcacgcccgattgcgagccttttcgacggcgacgcggggtcgc
841             --------+--------+--------+--------+--------+--------+        900
                ggggtacaaccttccgaagtgcgggctaacgctcggaaaagctgccgctgcgccccagcg

                gcggcaatttntccaacaaggtcagtggaccggcgcgccgatggccgcgcgcagccaggc
901             --------+--------+--------+--------+--------+--------+        960
                cgccgttaaanaggttgttccagtcacctggccgcgcggctaccggcgcgcgtcggtccg

                atccttggccggaaacacccgcgccgcatcatgatcggccaggatcgtccggcgcgcggc
961             --------+--------+--------+--------+--------+--------+       1020
                taggaaccggcctttgtgggcgcggcgtagtactagccggtcctagcaggccgcggcgcg

                gcggcgcaggtcggccgcgtcacccggattgtcaagcacccaggccatcgcgtccgcgac
1021            --------+--------+--------+--------+--------+--------+       1080
                cgccgcgtccagccggcgcagtgggcctaacagttcgtgggtccggtagcgcaggcgctg

                ctcgtccgcgtcgtccatgtcgacgatcaggccgttctccatgtcgcggaccagttcgcg
1081            --------+--------+--------+--------+--------+--------+       1140
                gagcaggcgcagcaggtacagctgctagtccggcaagaggtacagcgcctggtcaagcgc

                caccggggcggtgttcgatcgatcaccaggcatccggtggccatcgcctcggacagggac
1141            --------+--------+--------+--------+--------+--------+       1200
                gtggccccgccacaagctagctagtggtccgtaggccaccggtagcggagcctgtccctg

                caggaggtgacgaagggctcggtgaaatagacatgcgcgtgcgaggcctgcag
1201            --------+--------+--------+--------+--------+---    1253
                gtcctccactgcttcccgagccactttatctgtacgcgcacgctccggacgtc
```

FIG. 38B

```
      ATGAGACGAGACGTCAACCCGATCCACGCCACCCTTCTGCAGACCAGACTTGAGGAGATC
1     ---------+---------+---------+---------+---------+---------+   60
      TACTCTGCTCTGCAGTTGGGCTAGGTGCGGTGGGAAGACGTCTGGTCTGAACTCCTCTAG

      GCCCAGGGATTCGGTGCCGTGTCGCAGCCGCTCGGCCCGGCCATGAGCCATGGCGCGCTG
61    ---------+---------+---------+---------+---------+---------+  120
      CGGGTCCCTAAGCCACGGCACAGCGTCGGCGAGCCGGGCCGGTACTCGGTACCGCGCGAC

      TCGTCGGGCAAGCGTTTCCGCGGCATGCTGATGCTGCTTGCGGCAGAAGCCTCCCCCGGG
121   ---------+---------+---------+---------+---------+---------+  180
      AGCAGCCCGTTCGCAAAGGCGCCGTACGACTACGACGAACGCCGTCTTCGGAGCCCGCCC

      GTCTGCGACACGATCGTCGACGCCGCCTGCGCGGTCGAGATGGTGCATGCCGCATCGCTG
181   ---------+---------+---------+---------+---------+---------+  240
      CAGACGCTGTGCTAGCAGCTGCGGCGGACGCGCCAGCTCTACCACGTACGGCGTAGCGAC

      ATCTTCGACGACCTGCCCTGCATGGACGATGCCGGGCTGCGCCGCGGCCAGCCCGCGACC
241   ---------+---------+---------+---------+---------+---------+  300
      TAGAAGCTGCTGGACGGGACGTACCTGCTACGGCCCGACGCGGCGCCGGTCGGGCGCTGG

      CATGTGGCGCATGGCGAAAGCCGCGCCGTGCTAGGCGGCATCGCCCTGATCACCGAGGCG
301   ---------+---------+---------+---------+---------+---------+  360
      GTACACCGCGTACCGCTTTCGGCGCGGCACGATCCGCCGTAGCGGGACTAGTGGCTCCGC

      ATGGCCCTGCTGGCCGGTGCGCGCGGCGCGTCGGGCACGGTGCGGGCGCAGCTGGTGCGG
361   ---------+---------+---------+---------+---------+---------+  420
      TACCGGGACGACCGGCCACGCGCGCCGCGCAGCCCGTGCCACGCCCGCGTCGACCACGCC

      ATCCTGTCGCGGTCCCTGGGGCCGCAGGGCCTGTGCGCCGGCCAGGGCCTGGACCTGCAC
421   ---------+---------+---------+---------+---------+---------+  480
      TAGGACAGCGCCAGGGACCCCGGCGTCCCGGACACGCGGCCGGTCCTGGACCTGGACGTG

      GCGGCCAAGAACGGCGCGGGGGTCGAACAGGAACAGGACCTGAAGACCGGCGTGCTGTTC
481   ---------+---------+---------+---------+---------+---------+  540
      CGCCGGTTCTTGCCGCGCCCCCAGCTTGTCCTTGTCCTGGACTTCTGGCCGCACGACAAG

      ATCGCCGGGCTGGAGATGCTGGCCGTGATCAAGGAGTTCGACGCCGAGGAGCAGACTCAG
541   ---------+---------+---------+---------+---------+---------+  600
      TAGCGGCCCGACCTCTACGACCGGCACTAGTTCCTCAAGCTGCGGCTCCTCGTCTGAGTC

      ATGATCGACTTTGGCCGTCAGCTGGGCCGGGTGTTCCAGTCCTATGACGACCTGCTGGAC
601   ---------+---------+---------+---------+---------+---------+  660
      TACTAGCTGAAACCGGCAGTCGACCCGGCCCACAAGGTCAGGATACTGCTGGACGACCTG
```

FIG. 39A

```
      GTTGTGGGCGACCAGGCGGCGCTTGGCAAGGATACCGGTCGCGATGCGGCGGCCCCCGGC
661   ---------+---------+---------+---------+---------+---------+   720
      CAACACCCGCTGGTCCGCCGCGAACCGTTCCTATGGCCAGCGCTACGCCGCCGGGGGCCG

      CCGCGGCGCGGCCTTCTGGCCGTGTCAGACCTGCAGAACGTGTCCCGTCACTATGAGGCC
721   ---------+---------+---------+---------+---------+---------+   780
      GGCGCCGCGCCGGAAGACCGGCACAGTCTGGACGTCTTGCACAGGGCAGTGATACTCCGG

      AGCCGCGCCCAGCTGGACGCGATGCTGCGCAGCAAGCGCCTTCAGGCTCCGGAAATCGCG
781   ---------+---------+---------+---------+---------+---------+   840
      TCGGCGCGGGTCGACCTGCGCTACGACGCGTCGTTCGCGGAAGTCCGAGGCCTTTAGCGC

      GCCCTGCTGGAACGGGTTCTGCCCTACGCCGCGCGCGCCTAG
841   ---------+---------+---------+---------+--     882
      CGGGACGACCTTGCCCAAGACGGGATGCGGCGCGCGCGGATC
```

FIG. 39B

```
  1  MRRDVNPIHA TLLQTRLEEI AQGFGAVSQP LGPAMSHGAL SSGKRFRGML

 51  MLLAAEASGG VCDTIVDAAC AVEMVHAASL IFDDLPCMDD AGLRRGQPAT

101  HVAHGESRAV LGGIALITEA MALLAGARGA SGTVRAQLVR ILSRSLGPQG

151  LCAGQDLDLH AAKNGAGVEQ EQDLKTGVLF IAGLEMLAVI KEFDAEEQTQ

201  MIDFGRQLGR VFQSYDDLLD VVGDQAALGK DTGRDAAAPG PRRGLLAVSD

251  LQNVSRHYEA SRAQLDAMLR SKRLQAPEIA ALLERVLPYA ARA*
```

FIG. 40

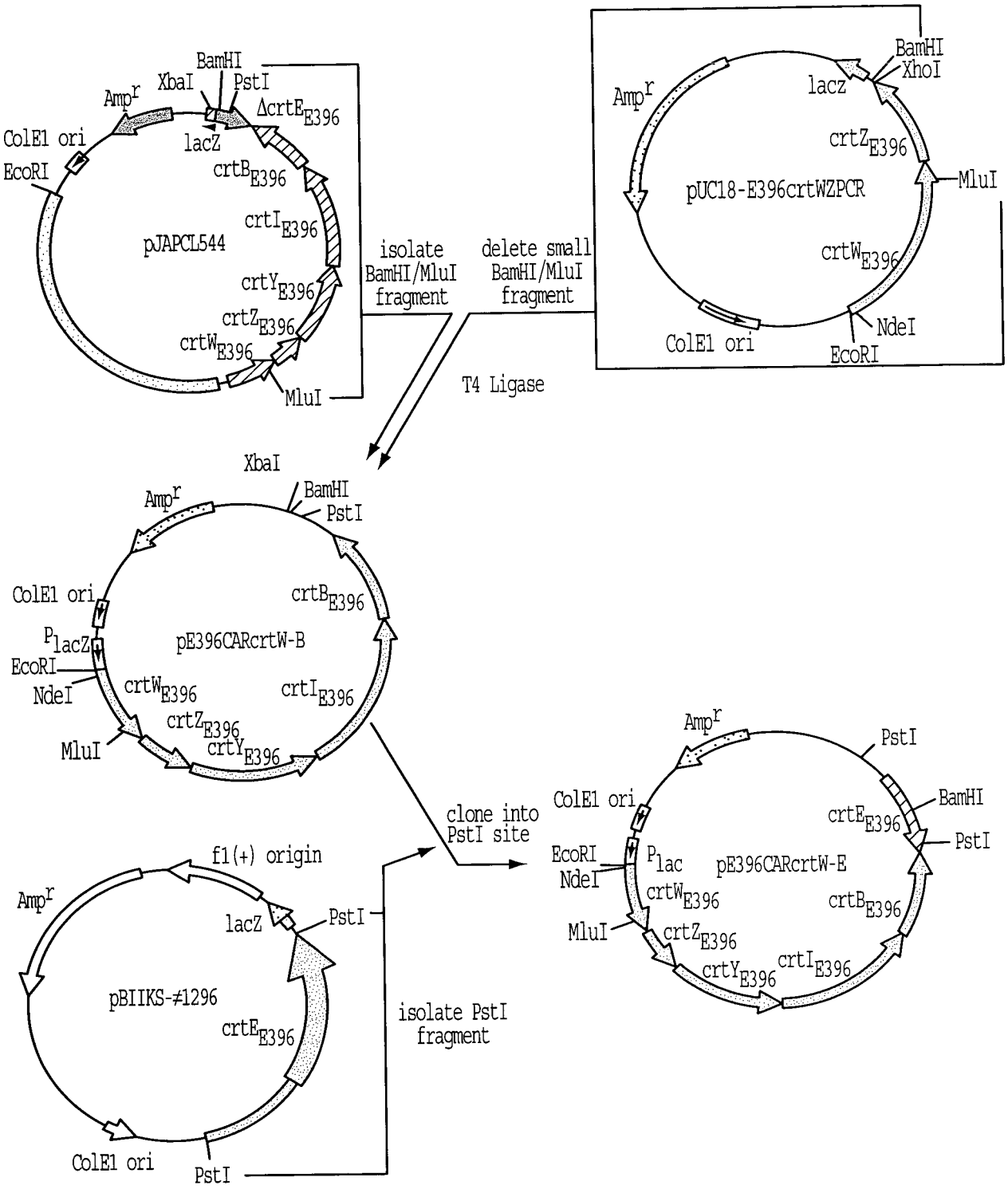

BamHI
XbaI
PstI
Amp^r
ΔcrtE_E396
ColE1 ori
lacZ
EcoRI
crtB_E396
crtI_E396
pJAPCL544
crtY_E396
crtZ_E396
crtW_E396
MluI
isolate BamHI/MluI fragment
delete small BamHI/MluI fragment
BamHI
XhoI
Amp^r
lacz
crtZ_E396
MluI
pUC18-E396crtWZPCR
crtW_E396
NdeI
ColE1 ori
EcoRI
T4 Ligase
XbaI
Amp^r
BamHI
PstI
ColE1 ori
crtB_E396
P_lacZ
pE396CARcrtW-B
EcoRI
NdeI
crtW_E396
crtI_E396
crtZ_E396
MluI
crtY_E396
clone into PstI site
Amp^r
PstI
ColE1 ori
BamHI
crtE_E396
PstI
EcoRI
NdeI
P_lac
pE396CARcrtW-E
crtW_E396
crtB_E396
MluI
crtZ_E396
crtY_E396
crtI_E396
f1(+) origin
Amp^r
lacZ
PstI
pBIIKS-≠1296
crtE_E396
isolate PstI fragment
ColE1 ori
PstI

FIG. 41

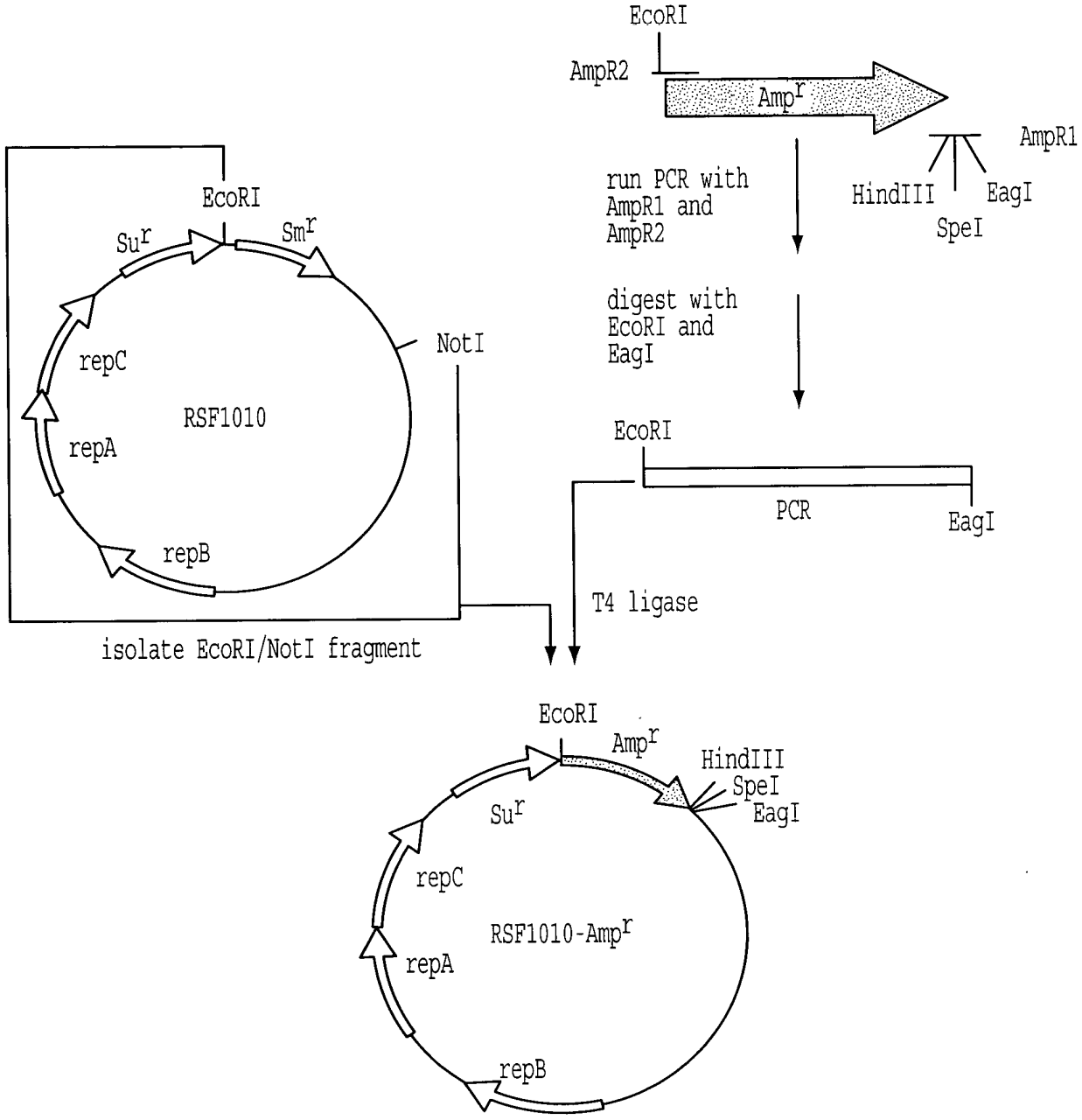

EcoRI
AmpR2
Amp^r
AmpR1
HindIII
EagI
SpeI
run PCR with AmpR1 and AmpR2
digest with EcoRI and EagI
EcoRI
PCR
EagI
T4 ligase
EcoRI
Su^r
Sm^r
NotI
repC
RSF1010
repA
repB
isolate EcoRI/NotI fragment
EcoRI
Amp^r
HindIII
SpeI
EagI
Su^r
repC
RSF1010-Amp^r
repA
repB

FIG. 42

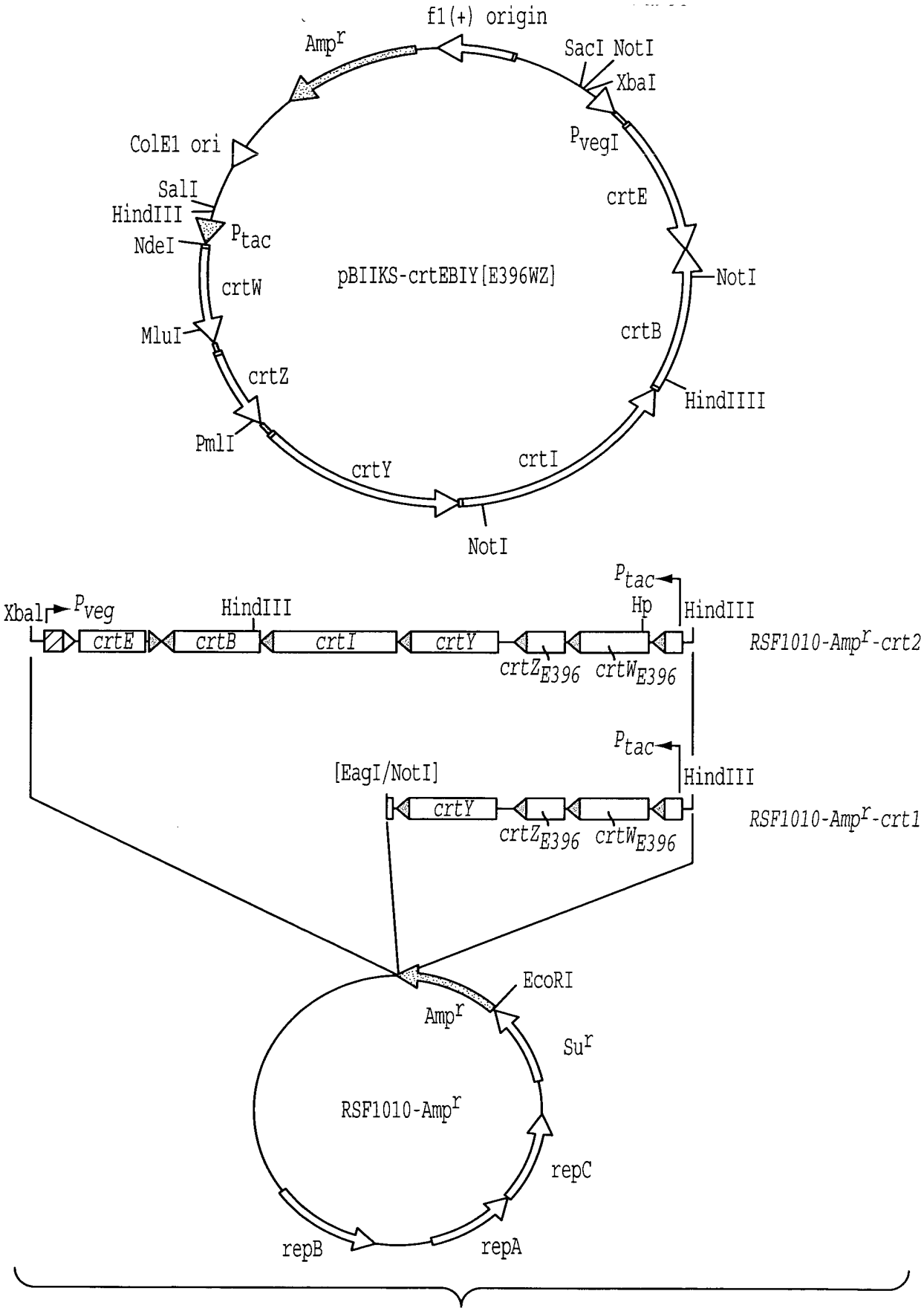

f1(+) origin
Amp^r
SacI NotI
XbaI
P$_{vegI}$
crtE
NotI
crtB
HindIIII
crtI
NotI
crtY
PmlI
crtZ
MluI
crtW
NdeI
P$_{tac}$
HindIII
SalI
ColE1 ori
pBIIKS-crtEBIY[E396WZ]
P$_{tac}$
Hp
HindIII
XbaI
P$_{veg}$
HindIII
crtE
crtB
crtI
crtY
crtZ$_{E396}$
crtW$_{E396}$
RSF1010-Amp^r-crt2
P$_{tac}$
[EagI/NotI]
HindIII
crtY
crtZ$_{E396}$
crtW$_{E396}$
RSF1010-Amp^r-crt1
EcoRI
Amp^r
Su^r
RSF1010-Amp^r
repC
repB
repA

FIG. 43